



**HAL**  
open science

# La tyrosyl-ARNt synthétase mitochondriale humaine : originalités fonctionnelles, structurales et place dans l'évolution

Luc Bonnefond

► **To cite this version:**

Luc Bonnefond. La tyrosyl-ARNt synthétase mitochondriale humaine : originalités fonctionnelles, structurales et place dans l'évolution. Biochimie [q-bio.BM]. Université Louis Pasteur - Strasbourg I, 2007. Français. NNT: . tel-00196315

**HAL Id: tel-00196315**

**<https://theses.hal.science/tel-00196315>**

Submitted on 12 Dec 2007

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ LOUIS PASTEUR  
STRASBOURG I

ÉCOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTÉ

## THÈSE

Présentée pour l'obtention du grade de

DOCTEUR DE L'UNIVERSITÉ LOUIS PASTEUR

Discipline : Sciences du vivant

Spécialité : Aspects Moléculaires et Cellulaires de la Biologie

Par

Luc Bonnefond

La tyrosyl-ARNt synthétase mitochondriale humaine :  
originalités fonctionnelles, structurales et place dans l'évolution

Soutenue le 22 Juin 2007 devant la commission d'examen

Prof. Sylvain Blanquet

Rapporteur externe

Dr. Chantal Abergel

Rapporteur externe

Dr. Anne-Catherine Dock-Brégeon

Rapporteur interne

Prof. Eric Westhof

Examineur

Dr. Joëlle Rudinger-Thirion

Directeur de thèse

Dr. Richard Giegé

Directeur de thèse



Je voudrais remercier toutes les personnes qui m'ont permis de réaliser ce travail depuis mon arrivée au laboratoire il à bientôt 5 ans jusqu'à son aboutissement aujourd'hui.

Tout d'abord, je souhaite remercier Mesdames Chantal Abergel, chargée de recherche à l'Institut de Biologie Structurale et Microbiologie (Luminy) et Anne-Catherine Dock-Brégeon, directrice de recherche à l'Institut de Génétique et de Biologie Moléculaire et Cellulaire (Strasbourg), ainsi que Messieurs Sylvain Blanquet, professeur à l'Ecole Polytechnique (Palaiseau), et Eric Westhof, professeur à l'Université Louis Pasteur pour l'honneur qu'ils me font en acceptant de juger ce travail.

Je remercie le Bernard Ehresmann pour m'avoir intéressé à la Biologie moléculaire puis chaleureusement ouvert les portes de son unité de recherche.

Je remercie également Catherine Florentz pour ses conseils et son soutien, tout au long de mon séjour dans son département.

Un très grand merci à Richard pour ses encouragements, son optimisme, ses conseils et les nombreuses discussions qui m'ont transmis sa soif de connaissances et son idéal d'excellence. Grâce à vous j'ai appris que la passion et la volonté permettent de surmonter de nombreux obstacles.

Toute ma gratitude va bien sûr à Joëlle. Merci pour ton soutien quotidien, même quand j'étais reclus dans les profondeurs de la cave, ta disponibilité, tes encouragements, ta générosité. Merci d'avoir cru en moi au DEA (et ce n'était pas gagné d'avance) et de m'avoir permis de m'épanouir au cours de ces 4 années et de découvrir de nouveaux horizons (Marrakech, la Californie,...). Ta "positive" attitude a permis de faire de cette thèse ce qui restera probablement l'une des plus belles périodes de ma vie.

Magali, mon 3<sup>ème</sup> directeur de thèse, merci pour tes conseils avisés, tes critiques toujours constructives (sauf peut-être en matière de musique) et ta disponibilité. Travailler sous tes ordres m'a donné cette fièvre du chercheur qui vous fait suivre des comptages d'aminoacylation en direct ou exploser une colonne entre les doigts...



Bernard, merci pour le temps que tu m'as consacré pour cette satanée enzyme, ton ingéniosité, tes conseils et l'intérêt que tu portes au travail des jeunes. Je n'oublierais pas ces fabuleuses nuits passées ensemble dans l'intimité des lignes du Synchrotron.

Marie, merci pour tes conseils, en matière de mitochondries notamment, mais aussi pour m'avoir fait participer au projet "Plaquette de l'UPR", à la Fête de la Science et m'avoir autorisé à descendre la climatisation à 25° en été.

Elodie, le rayon de soleil de la cristallographie, merci pour tout : la formation pratique accélérée en résolution de structures, le cinéma indépendant le dimanche soir, les mercredis de l'architecture, le théâtre nô, les concerts de grande musique, la randonnée et la tarte à la tomate.

Marie-Aline, merci les parties de squash acharnées et les longues discussions sur la condition du doctorant ou l'avenir incertain qui nous attend. Ton énergie et ta détermination m'ont aidé à venir à bout de cette thèse.

Merci à mes trois fées pour vos conseils avisés en matière d'image (lunettes, habillement) et en gestion de carrière. Merci pour vos encouragements, votre écoute. Grâce à vous je suis passé de l'ombre à la lumière (ou du noir aux couleurs).

Merci aussi à toute l'équipe du 443-447 pour ces 5 années ensemble dans cette ambiance joviale faite de discussions plus ou moins scientifiques autour de la HPLC ou d'un croissant.

Merci à tous ceux de l'IBMC qui ont contribué à faire de ces années une période agréable à tous points de vue et à toute heure. Spéciale dédicace à JC et Guillaume pour l'animation.

Merci à ma famille pour son soutien hebdomadaire, les 12 années à la Tour de Jade (merci mamie), les barbecues et la belotte estivale, Noël en trois fois ou avec le(s) chat(s) et la grasse matinée dominicale.

説子、いつもありがとう。一緒に住んでほしい、永遠に一緒に。愛してる。

# Abréviations

A	adénine
<i>A</i>	ampère
Å	angström (= 10 <sup>-10</sup> mètres)
aa	acide aminé
aaRS	aminoacyl-ARNt synthétase
ADN	acide désoxyribonucléique
ADNc	ADN copie
Amp	ampicilline
AMP	adénosine monophosphate
ARN	acide ribonucléique
ARNm	ARN messenger
ARNr	ARN ribosomique
ARNt	ARN de transfert
ATP	adénosine triphosphate
A <sub>x</sub>	absorbance à x nanomètres
BAP	phosphatase alcaline bovine
BET	bromure d'éthidium
β-ME	beta-mercaptoéthanol
BSA	albumine de sérum bovin
C	cytidine
Cam	chloramphénicol
CCD	détecteur à couplage de charge
Ci	curie
cm	centimètre
cpm	coups par minute
Da	dalton
DTE	dithioérythritol
DTT	dithiothréitol
DTNB	dithionitrobenzoate
DMSO	diméthylsulfoxyde
DO <sub>x</sub>	Densité optique à x nanomètres
EDTA	éthylène diamine tétraacétate de sodium
F	farad
g	gramme
<i>g</i>	gravité (unité d'accélération = 9,81 m.s <sup>-1</sup> )
G	guanine
GADPH	glyceraldéhyde 3-phosphate déshydrogénase
h	heure
HEPES	acide N-2 hydroxyl éthyl pipérazine N'2 éthane sulfoxyde
IPTG	isopropyl-beta-D-thiogalactopyranoside

k	kilo
Kan	kanamicine
$k_{cat}$	constante catalytique
$K_M$	constante de Michaelis
L	litre
m	mili
M	molaire
min	minute
mm	millimètre
mt	mitochondrial
$\mu$	micro
$\Omega$	ohm
n	nano
NiNTA	nickel nitrilotriacétate
NTP	nucléoside triphosphate
p	pico
PCR	réaction de polymérisation en chaîne
PEG	polyéthylène glycol
PPi	pyrophosphate
p/v	poids pour volume
SAP	phosphatase alcaline de crevette
SDS	sodium dodécyl sulfate
sec	seconde
SeMet	sélénométhionine
T	thymine
T4 PNK	polynucléotidokinase du phage T4
TBE	Tris-borate-EDTA
TCA	acide trichloroacétique
TEMED	N,N,N',N'-tetraméthyléthylènediamine
Tris	tris (hydroxyméthyl) amino éthane
U	uridine
$U$	unité
UV	ultra-violet
v	vitesse
V	volt
$V_{max}$	vitesse maximale
v/v	volume pour volume
X-Gal	5-bromo-4-chloro-3-indolyl-béta-D-galactopyranoside
$[S_0]$	concentration initiale en substrat
°C	degré Celsius

Chaque acide aminé est abrégé selon son code à trois lettres

Chaque aaRS est identifiée pour sa spécificité pour l'aa correspondant

# Table des matières

<b>Abréviations.....</b>	<b>5</b>
<b>Table des matières .....</b>	<b>7</b>
<b>Table des figures .....</b>	<b>11</b>
<b>Table des tableaux.....</b>	<b>11</b>
<b>Table des articles.....</b>	<b>12</b>
<b>INTRODUCTION .....</b>	<b>13</b>
<b>A. Les ARN de transfert .....</b>	<b>17</b>
1. Structure des ARNt .....	18
2. Biosynthèse et maturation des ARNt.....	19
3. Modifications post-transcriptionnelles des ARNt .....	20
4. Fonctions non canoniques des ARNt .....	21
5. Structures pseudo-ARNt .....	22
6. Evolution des ARNt .....	23
<b>B. Les aminoacyl-ARNt synthétases .....</b>	<b>25</b>
1. Les deux classes de synthétases.....	25
1.1. Les synthétases de classe I .....	26
1.2. Les synthétases de classe II .....	27
2. Structure et organisation modulaire des aminoacyl-ARNt synthétases .....	28
2.1. Organisation modulaire des aminoacyl-ARNt synthétases .....	28
2.2. Les complexes multisynthétasiques .....	32
2.3. Les aminoacyl-ARNt synthétases dupliquées .....	34
3. Evolution des aminoacyl-ARNt synthétases.....	34
4. Fonctions non-canoniques des aminoacyl-ARNt synthétases.....	35
4.1. Autorégulation de l'expression .....	35
4.2. Epissage d'introns.....	36
4.3. Implication dans le cycle viral .....	36
4.4. Molécules "signal".....	36
4.5. Voies métaboliques et régulation .....	37
4.6. Implications dans des pathologies.....	37
4.7. Les aminoacyl-ARNt synthétases comme cibles thérapeutiques.....	38
<b>C. Spécificité de la réaction d'amination.....</b>	<b>40</b>
1. Spécificité pour l'acide aminé .....	40
1.1. Édition des erreurs d'activation ou d'acylation d'acides aminés.....	40
1.2. Les domaines d'édition à l'état libre.....	41
2. Spécificité pour l'ARNt .....	42
3. Aminoacylation de mini substrats .....	44
4. Systèmes d'amination artificiels.....	45
5. Systèmes d'amination déviants .....	46
5.1. Les voies de formation indirectes des aminoacyl-ARNt .....	46
5.2. Les 21 <sup>ème</sup> et 22 <sup>ème</sup> acides aminés.....	47
<b>D. Les systèmes de tyrosylation .....</b>	<b>48</b>
Article de revue n°1	

<b>Évolution des systèmes d'aminocyclation ARN<sup>Tyr</sup>/TyrRS .....</b>	<b>48</b>
<b>E. Les aaRS mitochondriales humaines .....</b>	<b>63</b>
1. La mitochondrie .....	63
1.1. Structure et fonction .....	63
1.2. Génome .....	64
1.3. Synthèse protéique .....	65
1.4. Origine endosymbiotique .....	66
2. Les systèmes d'aminocyclation mitochondriaux humains .....	66
2.1. Les ARNt mitochondriaux .....	67
2.1.1. Originalités de séquence .....	67
2.1.2. Maturation .....	67
2.1.3. Structures "bizarres" .....	67
2.1.4. Implications dans des pathologies .....	68
2.2. Les aminoacyl-ARNt synthétases mitochondriales .....	69
2.3. Perte de spécificité des systèmes d'aminocyclation mitochondriaux .....	72
<b>F. Objectifs de la thèse .....</b>	<b>74</b>
<b>RÉSULTATS ET DISCUSSION .....</b>	<b>75</b>
<b>A. Obtention des molécules .....</b>	<b>77</b>
Article n°1	
<b>Évolution des systèmes d'aminocyclation spécifiques de la tyrosine .....</b>	<b>77</b>
Article n°2	
<b>Vers le jeu complet des aminoacyl-ARNt synthétases mitochondriales humaines : caractérisation de l'AspRS et de la TyrRS .....</b>	<b>83</b>
<b>B. Un jeu d'éléments d'identité restreint .....</b>	<b>96</b>
1. Perte de spécificité .....	96
Article n°3	
<b>La TyrRS mitochondriale humaine enfreint les règles de l'identité tyrosine .....</b>	<b>96</b>
2. Influence de la nature de la première paire de bases N1-N72 .....	104
3. Rôle accru de la base discriminatrice et de l'anticodon .....	105
Article de revue n°1	
<b>Évolution des systèmes d'aminocyclation ARN<sup>Tyr</sup>/TyrRS .....</b>	<b>105</b>
4. Mutations pathogéniques de l'ARNt <sup>Tyr</sup> mitochondrial .....	106
4.1. Effet des mutations sur l'aminocyclation par la TyrRS mitochondriale .....	107
4.2. Détermination de la structure des mutants pathogéniques du mt-ARNt <sup>Tyr</sup> mitochondrial humain .....	108
5. Le jeu d'identité est-il complet ? .....	112
<b>C. Caractérisation structurale de la TyrRS mitochondriale .....</b>	<b>113</b>
1. Obtention de cristaux de TyrRS mitochondriale exploitables pour la résolution de la structure. 113	
Article n°4	
<b>La tyrosyl-ARNt synthétase : premiers cristaux d'une aminoacyl-ARNt synthétase mitochondriale humaine .....</b>	<b>114</b>
2. La structure de la TyrRS mitochondriale .....	119
Article n°5	
<b>Propriétés générales et particularités de la structure de la TyrRS mitochondriale humaine .....</b>	<b>122</b>
3. Cystéines, ponts disulfure et activité .....	152
<b>D. Les TyrRS mitochondriales à travers l'évolution .....</b>	<b>155</b>
1. Recherche de séquences de TyrRS mitochondriales .....	155
2. Analyse phylogénétique .....	157

<b>CONCLUSIONS ET PERSPECTIVES .....</b>	<b>159</b>
Jeu d'identité du mt-ARNt <sup>Tyr</sup> et mt-TyrRS non spécifique .....	161
L'ARNt <sup>Tyr</sup> mitochondrial humain, jeu d'identité et structure .....	162
Structure de la mt-TyrRS, classique et singulière à la fois .....	161
La mt-TyrRS humaine, version longue, conformation et complexes .....	163
Les TyrRS, vers une vision globale .....	163
Les synthétases mitochondriales, nouvelles règles du jeu ? .....	165
<b>MATÉRIEL ET MÉTHODES.....</b>	<b>167</b>
<b>A. Matériel .....</b>	<b>169</b>
1. Produits chimiques .....	169
2. Enzymes commerciales .....	170
3. Enzymes non commerciales.....	170
4. Souches d' <i>Escherichia coli</i> .....	170
5. Plasmides .....	171
6. Milieux de culture .....	172
<b>B. Techniques courantes.....</b>	<b>173</b>
1. Préparation et transformation de bactéries compétentes .....	173
1.1. Méthode au CaCl <sub>2</sub> .....	173
1.2. Electrocompétence .....	173
2. Étude des acides nucléiques.....	174
2.1. Détermination de la concentration en acides nucléiques.....	174
2.2. Extraction phénolique .....	174
2.3. Précipitation des acides nucléiques.....	174
2.4. Amplification d'ADN plasmidique .....	175
2.5. Migration des acides nucléiques sur gels.....	175
2.6. Éluion des ARN à partir d'un gel de polyacrylamide.....	176
3. Étude des protéines .....	176
3.1. Détermination de la concentration en protéines.....	176
3.2. Électrophorèse en conditions dénaturantes .....	177
3.4. Diffusion de lumière.....	177
<b>C. Préparation d'ARNt par clonage et transcription <i>in vitro</i>.....</b>	<b>180</b>
1. Principe.....	180
2. Construction des gènes synthétiques .....	180
3. Clonage .....	181
4. Sélection des clones positifs.....	181
5. Transcription <i>in vitro</i> .....	182
6. Purification des transcrits sur gels préparatifs .....	182
7. Analyse de l'extrémité 5' des ARNt transcrits .....	183
7.1. Principe .....	183
7.2. Marquage de l'extrémité 3' des transcrits au [ $\gamma$ - <sup>32</sup> P]-pCp .....	183
7.3. Hydrolyse des transcrits .....	183
7.4. Chromatographie bidimensionnelle.....	183
7.5. Révélation.....	184
<b>D. Clonage, surexpression et purification de la TyrRS mitochondriale humaine et de ses variants.....</b>	<b>186</b>
1. Clonage du gène à partir d'une banque d'ADNc humains .....	186
1.1. Amplification du gène par PCR.....	186
1.2. Clonage dans le vecteur pCR2.1.....	186
1.3. Sous-clonage dans le vecteur pQE70.....	187

1.3.1. Purification de l'insert .....	187
1.3.2. Digestion et déphosphorylation du vecteur PQE70.....	187
1.3.3. Ligation .....	187
1.3.4. Transformation des cellules TG1 électrocompétentes.....	188
2. Clonage des variants .....	190
2.1. La mt-TyrRS tronquée et le domaine C-terminal .....	190
2.2. Mutagenèse dirigée .....	190
3. Expression.....	190
4. Purification sur colonnes.....	193
4.1. Colonne d'affinité NiNTA.....	193
4.2. Colonne d'exclusion de taille.....	194
5. Dialyses .....	194
<b>E. Tests fonctionnels.....</b>	<b>195</b>
1. Aminoacylation <i>in vitro</i> d'ARNt .....	195
1.1. Principe du test d'aminoacylation .....	195
1.2. Les paramètres cinétiques étudiés.....	195
1.3. Le test d'aminoacylation.....	196
2. Echange ATP/[ <sup>32</sup> P]-PPi.....	197
<b>F. Cartographie des ARNt en solution .....</b>	<b>198</b>
1. Principe.....	198
2. Marquage des ARNt en 5' .....	198
3. Cartographie .....	199
<b>G. Cristallographie.....</b>	<b>200</b>
1. Principe de la cristallogénèse .....	200
2. Cristallisation par diffusion de vapeur.....	201
3. Diffraction des rayons X.....	201
4. Résolution de structure.....	204
<b>H. Bioinformatique .....</b>	<b>205</b>
1. Alignements de séquences.....	205
2. Arbres phylogénétiques .....	205
3. Outils bioinformatiques.....	205
<b>Annexes .....</b>	<b>207</b>
Articles.....	209
Présentations orales .....	209
Affiches.....	210
Séquences des TyrRS mitochondriales.....	211
Résumé en français.....	222
Résumé en anglais .....	223
<b>RÉFÉRENCES BIBLIOGRAPHIQUES.....</b>	<b>225</b>

# Table des figures

<b>Figure 1.</b> Structures secondaires et tertiaires des ARNt canoniques	17
<b>Figure 2.</b> Structures pseudo-ARNt	22
<b>Figure 3.</b> Etapes évolutives de l'ARNt	23
<b>Figure 4.</b> Superposition des domaines catalytiques des aaRS de classes I et II	29
<b>Figure 5.</b> Architecture modulaire des aminoacyl-ARNt synthétases de classe I	30
<b>Figure 6.</b> Architecture modulaire des aminoacyl-ARNt synthétases de classe II	31
<b>Figure 7.</b> Le complexe multisynthétasique de mammifères	33
<b>Figure 8.</b> Localisation et importance des nucléotides des jeux d'identité des ARNt d' <i>E. coli</i>	43
<b>Figure 9.</b> Structures aminoacylables minimalistes dérivées du bras accepteur d'ARNt	44
<b>Figure 10.</b> Arbre phylogénétique des TyrRS	62
<b>Figure 11.</b> Organisation schématique de la mitochondrie et de ses constituants	63
<b>Figure 12.</b> Organisation du génome mitochondrial humain	65
<b>Figure 13.</b> Affiche pour le "20 <sup>th</sup> International tRNA Workshop" (Banz, Allemagne, 2-7 Octobre 2003)	82
<b>Figure 14.</b> Affiche pour le "Congrès annuel de la SFBBM" et le "32 <sup>ème</sup> Forum des Jeunes Chercheurs" (Nantes, 24-26 Octobre 2005)	102
<b>Figure 15.</b> Affiche pour le "21 <sup>st</sup> International tRNA Workshop", (Bangalore, Inde, 2-7 Décembre 2005)	103
<b>Figure 16.</b> Mutations pathologiques de l'ARNt <sup>Tyr</sup> mitochondrial humain	107
<b>Figure 17.</b> Cartographie en solution de transcrits de l'ARNt <sup>Tyr</sup> mitochondrial humain	109
<b>Figure 18.</b> Prédictions de structures secondaires de l'ARNt <sup>Tyr</sup> mitochondrial humain	110
<b>Figure 19.</b> Séquences et structures secondaires d'ARNt <sup>Tyr</sup> mitochondriaux	111
<b>Figure 20.</b> Formules des dérivés d'adénylate de tyrosine utilisés en cristallographie	119
<b>Figure 21.</b> Structure de la mt-TyrRS- $\Delta$ S4 complexée à l'analogue de tyrosyl-adénylate TYA	120
<b>Figure 22.</b> Photos de cristaux de mt-TyrRS- $\Delta$ S4 marquée au sélénium	121
<b>Figure 23.</b> Implication des ponts disulfures dans la conformation de la mt-TyrRS- $\Delta$ S4	154
<b>Figure 24.</b> Arbre phylogénétique des TyrRS mitochondriales.	158
<b>Figure 25.</b> Superposition des structures de TyrRS eubactériennes avec domaine S4-like	164
<b>Figure 26.</b> Production de "transzymes"	179
<b>Figure 27.</b> Principe du clonage de la mt-TyrRS humaine dans le vecteur d'expression pQE70	185
<b>Figure 28.</b> Principe de la mutagenèse dirigée selon Stratagene	189
<b>Figure 29.</b> Purification par affinité de la mt-TyrRS- $\Delta$ S4 sur colonne NiNTA	192
<b>Figure 30.</b> Diagramme de solubilité théorique	200
<b>Figure 31.</b> Dispositif expérimental utilisé pour l'enregistrement de clichés de diffraction	203



## Table de tableaux

<b>Tableau 1.</b> Principales caractéristiques des aaRS et répartition en classes et sous-classes	26
<b>Tableau 2.</b> Structures cristallographiques de TyrRS connues déposées à la PDB	61
<b>Tableau 3.</b> Caractéristiques des aminoacyl-ARNt synthétases mitochondriales humaines	71
<b>Tableau 4.</b> Paramètres cinétiques de l'aminoylation des aaRS mitochondriales humaines	70
<b>Tableau 5.</b> Paramètres cinétiques de l'aminoylation de transcrits de l'ARNt <sup>Tyr</sup> mitochondrial muté en N1-N72	104
<b>Tableau 6.</b> Paramètres cinétiques de l'aminoylation de transcrits de l'ARNt <sup>Tyr</sup> mitochondrial mutés au niveau du triplet anticodon et de la base discriminatrice	106
<b>Tableau 7.</b> Paramètres cinétiques de l'aminoylation de transcrits de l'ARNt <sup>Tyr</sup> mitochondrial mutés à des positions pathogéniques	108
<b>Tableau 8.</b> Jeux de données pour les formes libre et complexées de la mt-TyrRS- $\Delta$ S4	120
<b>Tableau 9.</b> Séquences de TyrRS mitochondriales	156

## Table des articles

Article de revue n°1

<b>Évolution des systèmes d'aminoylation ARNt<sup>Tyr</sup>/TyrRS</b> .....	48
---	----

Article n°1

<b>Evolution des systèmes d'aminoylation spécifiques de la tyrosine</b> .....	77
---	----

Article n°2

<b>Vers le jeu complet des aminoacyl-ARNt synthétases mitochondriales humaines : caractérisation de l'AspRS et de la TyrRS</b> .....	83
--	----

Article n°3

<b>La TyrRS mitochondriale humaine enfreint les règles de l'identité tyrosine</b> .....	96
---	----

Article n°4

<b>La tyrosyl-ARNt synthétase : premiers cristaux d'une aminoacyl-ARNt synthétase mitochondriale humaine</b> .....	114
--	-----

Article n°5

<b>Propriétés générales et particularités de la structure de la TyrRS mitochondriale humaine</b> .....	122
--	-----

# Introduction



La traduction de l'information génétique est le processus essentiel permettant l'expression de l'information génétique portée par l'ADN. Elle a lieu au niveau des ribosomes où les codons de l'ARN messager (ARNm) du gène exprimé sont lus et traduits en acides aminés pour former une protéine. Le code génétique définit les règles d'équivalence codon/acide aminé et l'ARN de transfert (ARNt) tient lieu d'interprète au niveau du ribosome. Il assure la correspondance entre le codon de l'ARNm et l'acide aminé fixé sur l'ARNt par la complémentarité codon/anticodon. Ainsi, la fidélité de la traduction repose sur cette association, mais aussi sur celle préalable de l'ARNt avec l'acide aminé. Cette dernière association, appelée réaction d'aminoacylation, consiste en la fixation covalente d'un acide aminé sur l'extrémité 3' de l'ARNt correspondant. Cette réaction est catalysée par une classe d'enzymes appelées aminoacyl-ARNt synthétases (aaRS). Chaque cellule est dotée d'une vingtaine d'aaRS, chacune devant être spécifique d'un acide aminé et d'une famille d'ARNt isoaccepteurs. La spécificité est assurée par des éléments présents sur l'aaRS et l'ARNt. L'ensemble de ces éléments constitue le "second code génétique" (Lengyel, 1966; de Duve, 1988). Ce "code", dans ses grandes lignes, a été déchiffré chez les bactéries et dans une moindre mesure chez les archaea et les eucaryotes.

La molécule d'ARNt interagit avec de nombreuses protéines cellulaires en plus des aaRS. Mais à la différence des synthétases, les autres interactants tels que le facteur d'élongation EF-Tu, la RNase P, la CCase ou les enzymes de maturation interagissent indistinctement avec tous les ARNt en reconnaissant des éléments communs. Les ARNt présentent donc à la fois (i) des signaux consensus, (ii) une architecture commune répondant à des contraintes structurales, pour l'interaction avec le ribosome notamment, et (iii) des éléments permettant de les distinguer les uns des autres. Le cumul de ces différentes contraintes a abouti à la structure secondaire des ARNt dite en "feuille de trèfle".

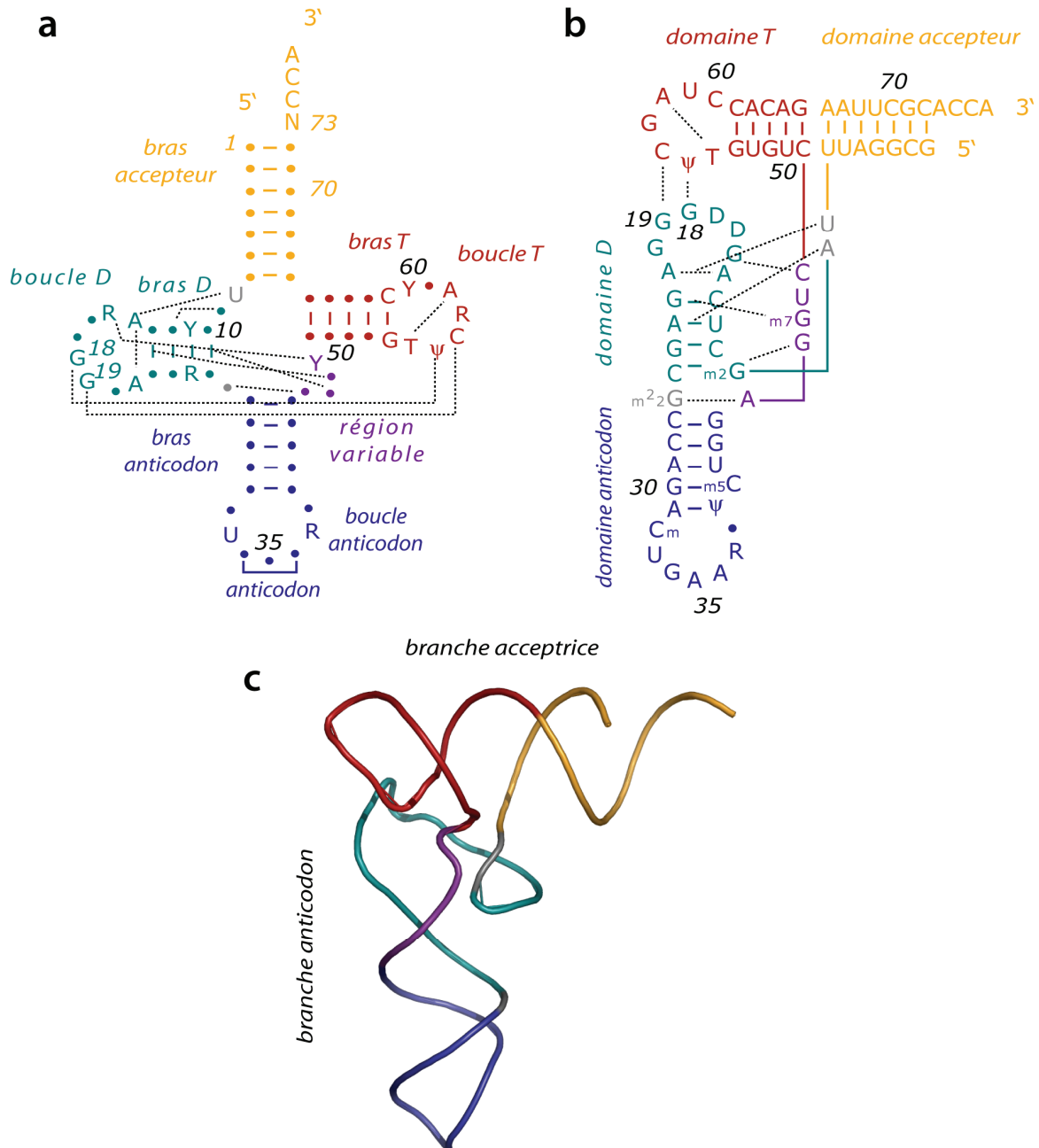
Mon travail de thèse est focalisé sur les systèmes d'aminoacylation. Parmi eux, celui de la tyrosine a été largement étudié ces 40 dernières années. La tyrosyl-ARNt synthétase (TyrRS) de *Bacillus stearothermophilus* est d'ailleurs la première aaRS dont la structure cristallographique a été résolue et a fait l'objet d'études enzymologiques très poussées dans les années 70 (pour revue Fersht, 1987). Cependant, de nombreuses questions restent sans réponse aussi bien au niveau structural que fonctionnel et évolutif. De plus, des différences fonctionnelles significatives dans le mécanisme de reconnaissance entre l'ARNt<sup>Tyr</sup> et la TyrRS ont été décrites entre les bactéries et les archaea/eucaryotes. En effet, contrairement à ce qui a été observé pour les autres systèmes d'aminoacylation, les réactions de charges croisées entre ces deux groupes sont impossibles. Ceci est dû d'une part à la nature différente de la première paire de bases 1-72 de l'ARNt<sup>Tyr</sup> et d'autre part à l'organisation

particulière du site de fixation de la branche acceptrice de l'ARNt<sup>Tyr</sup> des TyrRS bactériennes et archaea/eucaryotiques.

Ma thèse porte sur l'étude fonctionnelle et structurale des partenaires de la réaction d'aminacylation spécifique de la tyrosine dans la mitochondrie humaine. En effet, l'origine endosymbiotique et l'évolution de la mitochondrie en font un système hybride entre eubactérie et eucaryote.

L'introduction de cette thèse se divise en 5 parties. Les deux partenaires de la réaction d'aminacylation, l'ARN de transfert (ARNt) et l'aminacyl-ARNt synthétase (aaRS), seront présentés successivement, sous leurs aspects fonctionnels, structuraux, et évolutifs. Tout en soulignant le rôle primordial de l'ARNt et de l'aaRS dans la réaction d'aminacylation, leurs implications dans d'autres processus cellulaires seront développées. La troisième partie se focalisera sur la réaction d'aminacylation et sa spécificité, mais aussi sur les déviations aux règles générales établies. L'attention sera ensuite portée sur le système d'aminacylation spécifique de la tyrosine au travers d'un article de revue. Les particularités intrinsèques de ce système et ses différences entre phylae seront détaillées. Le cinquième volet sera dédié aux systèmes d'aminacylation mitochondriaux humains. Après avoir rappelé le contexte cellulaire, génétique et évolutif de la mitochondrie, l'état des connaissances sur les ARNt et aaRS mitochondriaux sera établi. Les originalités de ces systèmes d'aminacylation seront mises en exergue.

# A. Les ARN de transfert



**Figure 1. Structures secondaires et tertiaires des ARNt canoniques**

Les différents domaines structuraux sont indiqués. Les interactions à longues distances permettant le repliement tridimensionnel sont représentées par des traits en pointillés. **(a)** Structure secondaire consensus en feuille de trèfle des ARNt canoniques. **(b)** Structure secondaire de l'ARNt<sup>Phe</sup> de *Saccharomyces cerevisiae* correspondant à sa structure cristallographique **(c)** (Jovine et al., 2000). La numérotation des nucléotides ainsi que la nomenclature des bases modifiées correspondent à celles établies par Sprinzl et Vassilenko (2005). Y et R indiquent des pyrimidines et des purines, respectivement.

Les ARN de transfert (ARNt) sont des ARN non codants qui jouent le rôle d'adaptateurs moléculaires pour traduire le langage nucléaire de l'information génétique en langage protéique. Leur existence avait été postulée dès 1955 par Crick (Crick, 1955) puis confirmée expérimentalement trois ans plus tard (Hoagland et al., 1958). L'ARNt<sup>Ala</sup> de *S. cerevisiae* fut le premier séquencé (Holley et al., 1965) et l'ARNt<sup>Phe</sup> le premier dont la structure cristallographique fut résolue (Kim et al., 1974 ; Robertus et al., 1974). Les nombreuses études sur les ARNt ont permis de déterminer leurs caractéristiques structurales, de préciser leur fonction et leur évolution. Le site de compilation de séquences et de gènes d'ARNt fait ainsi état de plus de 5800 séquences connues provenant de 111 organismes différents (Sprinzl et Vassilenko, 2005; <http://www.tRNA.uni-bayreuth.de>). Parallèlement, plus de 90 types de modifications post-transcriptionnelles ont été identifiées au cours du séquençage de ces ARNt d'après la banque de données de modifications d'ARN (Limbach et al., 1994 ; <http://library.med.utah.edu/RNAmods/>).

## 1. Structure des ARNt

Les séquences d'ARNt connues comportent entre 75 et 95 nucléotides, parmi lesquels 15 sont strictement conservés chez tous les ARNt et 8 semi-conservés (**figure 1.a**). Les complémentarités de séquence permettent un repliement secondaire dit en "feuille de trèfle" qui s'applique à l'ensemble des ARNt. Cette structure comprend 4 bras ou hélices, constituées essentiellement de paires de bases Watson-Crick, reliées entre elles par des régions de un à deux nucléotides (à l'exception de la région variable) et fermées pour trois d'entre elles par des boucles (boucles D, anticodon et T). On divise généralement l'ARNt en 5 domaines : le domaine accepteur, le domaine D, le domaine anticodon, la région variable et le domaine T (**figure 1.a**). Le domaine accepteur comporte généralement 7 paires de bases et est toujours surmonté de 4 nucléotides non appariés : la base discriminatrice en position 73 et la séquence conservée CCA 3' terminale, site de fixation de l'acide aminé. Les bras anticodon et T possèdent 5 paires de bases contre 4 pour le bras D. La boucle du bras anticodon est formée de 7 nucléotides dont le triplet central 34-35-36 constitue l'anticodon. La boucle du bras T est aussi formée de 7 nucléotides, mais la boucle du bras D comporte entre 7 et 12 nucléotides. Enfin, la région variable présente la plus grande diversité de taille et compte de 4 à 23 nucléotides selon l'ARNt concerné. La taille de cette région a permis de diviser les ARNt en deux groupes (Rich et RajBhandary, 1976). Le premier groupe englobe les ARNt avec une petite région variable de 4 à 5 nucléotides, alors que le second regroupe

les ARNt à grande région variable, laquelle peut alors adopter une structure en tige boucle. Seuls 4 ARNt appartiennent à ce deuxième groupe, il s'agit de l'ARNt<sup>Tyr</sup> de type bactérien, l'ARNt<sup>Leu</sup>, l'ARNt<sup>Ser</sup> et l'ARNt<sup>Sec</sup>. Il est à noter que certains ARNt dérogent au repliement standard comme l'ARNt<sup>His</sup> eubactérien qui contient 8 paires de bases dans le bras accepteur; la 8<sup>ème</sup> paire de bases est formée par l'appariement entre un résidu supplémentaire en 5' et la base 73. Les ARNt adoptent une structure tridimensionnelle en forme de "L" grâce à un jeu d'interactions entre les nucléotides conservés des boucles D et T (**figure 1.b et 1.c**). Les deux branches du "L" correspondent à deux hélices perpendiculaires formées d'une part par l'empilement des bras accepteur et T et d'autre part par l'empilement des bras D et anticodon. Cette structure en "L" peut présenter des différences notamment dans l'angle formé par les deux hélices qui varie entre 90 et 110°, mais aussi par la présence d'un réseau alternatif d'interactions tertiaires (Baron et al., 1993), ou par l'absence de domaines (Kumazawa et al., 1988). De plus, la structure de l'ARNt n'est pas figée, mais évolue de façon dynamique lors de la reconnaissance par ses différents partenaires (par exemple, Ishitani et al., 2003).

## 2. Biosynthèse et maturation des ARNt

Les gènes d'ARNt sont organisés différemment dans les génomes des bactéries et des eucaryotes. Chez les bactéries, ils sont généralement organisés en opérons avec d'autres gènes d'ARNt, d'ARNr ou de protéines, et leurs promoteurs sont constitués des éléments -35 et -10. Chez les eucaryotes, les gènes d'ARNt sont isolés les uns des autres et leurs éléments promoteurs pour la transcription par la polymérase III sont les boîtes A et B situées dans la séquence du gène (positions 8 à 19 pour la boîte A et 52 à 62 pour la boîte B). Les gènes d'ARNt sont transcrits sous la forme de pré-ARNt comportant des extensions 5' et 3', la séquence CCA chez les bactéries et des introns chez certains eucaryotes et archaea. La maturation en 5' est commune à tous les ARNt et fait intervenir une endonucléase, la ribonucléase P (RNase P) qui est une ribonucléoprotéine dont l'activité catalytique est portée par l'acide nucléique (James et al., 1988). A l'opposé, la maturation en 3' fait intervenir de nombreuses enzymes et son mécanisme diffère entre les trois grands phylae. Chez les bactéries une endonucléase réalise une hydrolyse partielle puis une combinaison d'exonucléases libère une extrémité CCA intacte. Chez les eucaryotes et les archaea, l'action des endo- et exonucléases libère le gène de l'ARNt dépourvu de la séquence CCA 3' terminale. Celle-ci est ajoutée par l'(ATP, CTP) : ARNt nucléotidyl transférase (Deutscher,



1983) et les éventuels introns (principalement présents après le nucléotide 37 de la boucle anticodon) sont excisés (Knapp et al., 1979 ; Xue et al., 2006). De façon surprenante, l'archaea *Nanoaquifex equitans* comporte dans son génome 18 mini-gènes de demi-ARNt qui permettent de constituer 9 ARNt fonctionnels après transcription et maturation (Randau et al., 2005).

### 3. Modifications post-transcriptionnelles des ARNt

A ce point du processus de maturation, les ARNt ne sont pas encore fonctionnels. Des modifications spécifiques de leurs nucléotides sont encore nécessaires pour leur conférer leurs propriétés fonctionnelles et structurales définitives. Les ARNt présentent une très grande diversité de modifications (plus de 90 ont été répertoriées : <http://library.med.utah.edu/RNAmods/>) qui peuvent être spécifiques des ARNt (par exemple la queuosine, qui est un dérivé de guanine présent dans l'anticodon de certains ARNt<sup>Asp</sup> et ARNt<sup>Tyr</sup> notamment) ou retrouvées chez d'autres types d'ARN. Elles sont plus nombreuses chez les eucaryotes que chez les bactéries. Tandis que certaines sont communes à tous les ARNt (telles que la pseudouridine de la boucle T), d'autres sont spécifiques d'un ARNt (par exemple la wybutosine, dérivé de la guanine, spécifique des ARNt<sup>Phe</sup> d'eucaryotes et d'archaea ; Blobstein et al., 1973) et finalement certaines sont présentes uniquement dans un domaine du vivant (par exemple l'archaeosine, un dérivé de guanine, spécifique des archaea ; Gregson et al., 1993). Elles contribuent (i) à la structuration de l'ARNt (Perret et al., 1990) jusqu'à empêcher un repliement incorrect comme le m<sup>1</sup>A9 de l'ARNt<sup>Lys</sup> mitochondrial qui est nécessaire à la formation de la feuille de trèfle (Helm et al., 1999). (ii) Elles protègent l'ARNt contre les dégradations (Marinus et al., 1975) et lui confèrent une plus grande thermostabilité (Horie et al., 1985). (iii) Elles peuvent aussi être directement impliquées dans la reconnaissance en tant qu'éléments d'identité spécifiquement reconnus par l'aaRS homologue (telles que l'inosine en position "wobble" de l'ARNt<sup>Ile</sup> de *S. cerevisiae* ; Senger et al., 1997) ou jouer le rôle d'antidéterminant en empêchant la reconnaissance d'un ARNt par une aaRS hétérologue (citons la lysidine en position 34 de l'ARNt<sup>Ile</sup> de *E. coli* (Muramatsu et al., 1988) ou la modification m<sup>1</sup>G37 de l'ARNt<sup>Asp</sup> de *S. cerevisiae* (Perret et al., 1990)). Elles jouent aussi un rôle essentiel lors de la reconnaissance codon/anticodon où la modification du nucléotide en position "wobble" de l'anticodon permet de reconnaître les différents codons spécifiques d'un même acide aminé assurant ainsi la dégénérescence du code génétique. Toutes ces modifications sont réalisées par des enzymes spécifiques ayant

fait l'objet d'études fonctionnelles et structurales (description des voies de modifications d'ARN connues sur le site MODOMICS et revue des structures d'enzymes de modification connues par Grosjean et Benne, 1998 ; Nakanishi et Nureki, 2005).

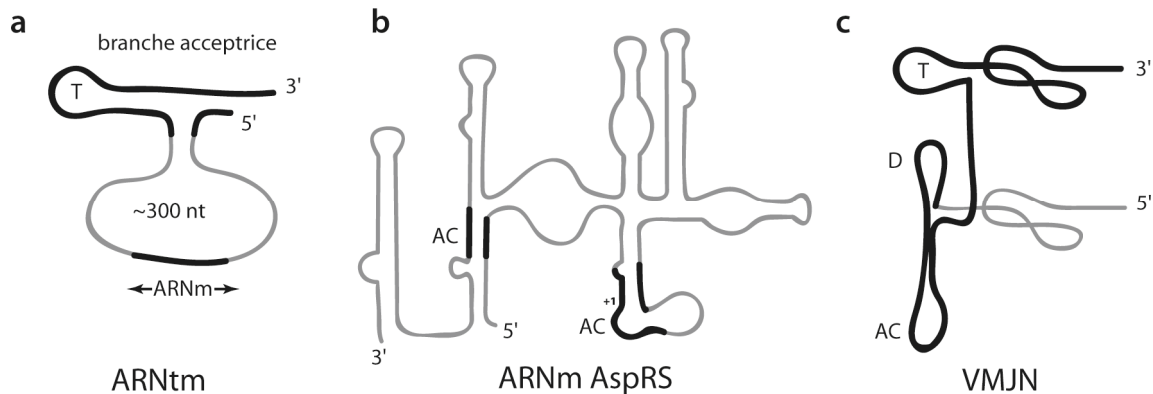
#### 4. Fonctions non canoniques des ARNt

La présence des ARNt en quantité adéquate dans les cellules permet une synthèse protéique efficace et facilite leur participation à d'autres processus cellulaires (Söll, 1993). Parmi les nombreuses fonctions que peuvent assurer ces ARNt, nous ne citerons que quelques exemples parmi les plus marquants.

Les ARNt peuvent participer à la régulation du taux d'acide aminé libre. Ainsi, chez *Bacillus subtilis* et chez certaines bactéries gram positif, l'ARNt libre se fixe sur l'ARNm de son aaRS spécifique au niveau d'une région régulatrice. Cette région de 300 nucléotides est composée d'une série de tiges-boucles et inclut une séquence conservée appelée "T box" ainsi qu'une tige-boucle terminatrice. La fixation de l'ARNt par son anticodon et son extrémité CCA-3' libre au niveau de cette région régulatrice induit un changement de conformation de la structure terminatrice en structure anti-terminatrice (Putzer et al., 1995). A un autre niveau, plusieurs ARNt sont détournés de la machinerie cellulaire où ils servent d'amorces à la rétro-transcription de virus. Cette action fait intervenir la complémentarité de séquence entre l'extrémité 3' de l'ARNt et un site dit PBS (Primer Binding Site) dans la partie 5' du génome viral (Marquet et al., 1995). Ce phénomène concerne principalement l'ARNt<sup>Lys3</sup> pour les virus tels que le VIH (Virus de l'Immunodéficience Humaine), l'ARNt<sup>Pro</sup> pour le virus de Moloney, et l'ARNt<sup>Trp</sup> pour le virus de la leucose aviaire. Les rétrotransposons à LTR (Long Term Repeat) utilisent aussi des ARNt comme amorce lors de la rétro-transcription. Dans ce cas, leur PBS est complémentaire d'une séquence interne de l'ARNt cible. Enfin, les rétrotransposons Ty1 et Ty2 de *S. cerevisiae* utilisent l'ARNt<sub>i</sub><sup>Met</sup> comme amorce. On retrouve aussi des ARNt impliqués dans des voies métaboliques. Ainsi l'ARNt<sup>Glu</sup> chloroplastique préalablement aminoacylé est réduit par la glutamyl-ARNt réductase dans la première étape de synthèse de tétrapyrolle chez la plupart des bactéries, archaea et dans les chloroplastes. Cette voie de synthèse aboutit notamment à la forme de l'hème et de la chlorophylle (Schön et al., 1986). Citons aussi l'ARNt<sup>Gly</sup> qui est impliqué dans le mécanisme de synthèse du peptidoglycane de la paroi bactérienne de *Staphylococcus aureus* (Kamiryo et Matsushashi, 1969 ; Rohrer et Berger-Bachi, 2003).

## 5. Structures pseudo-ARNt

L'ARNt est une molécule complexe, fruit d'une longue évolution (**cf. A.6.**), dont la fonction est centrale lors de la traduction où il interagit avec de nombreux partenaires. Les intermédiaires et variants générés au cours de ce processus évolutif ont été mis à profit par la nature pour participer à des processus biologiques autres que la traduction.



**Figure 2. Structures pseudo-ARNt**

Les mimes de domaines d'ARNt sont indiqués en noir (T pour domaine T, D pour domaine D et AC pour domaine anticodon). **(a)** Représentation schématique de l'ARNtm (transfert-messager). **(b)** Partie 5' de l'ARNm de l'AspRS de *S. cerevisiae* mimant deux branches anticodon. **(c)** Pseudo-ARNt retrouvé en 3' de l'ARN génomique du virus de la mosaïque jaune du navet (VMJN). Figure adaptée de Giegé et Frugier (2003) et Ryckelynck et al. (2005).

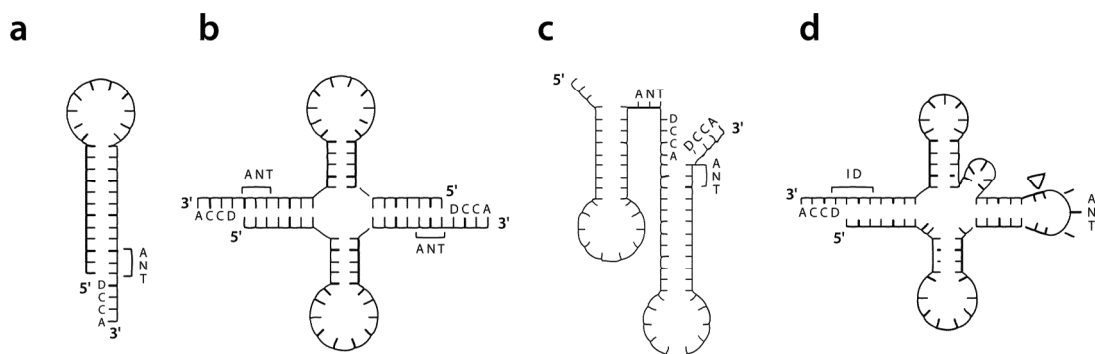
L'ARNtm de *E. coli* (pour transfert message) est impliqué dans une voie de dégradation des polypeptides issus d'ARNm dépourvus de codons stop et dans le recyclage des ribosomes bloqués sur ses ARNm (Keiler et al., 1996). L'ARNtm joue, comme son nom l'indique, à la fois le rôle d'ARN de transfert et de messenger. En effet, il est aminoacylé par l'AlaRS à son extrémité 3' grâce au mime structural de la branche acceptrice de l'ARNt<sup>Ala</sup> (**figure 2.a.**). D'autre part il possède une séquence codant pour 10 acides aminés qui, accrochés à l'extrémité C-terminale du polypeptide, serviront de signal de reconnaissance pour sa dégradation par différentes protéases. C'est la structure pseudo-ARNt qui permet aussi la reconnaissance de cette molécule par le facteur d'élongation EF-Tu qui l'achemine jusqu'au ribosome.

On retrouve également des structures pseudo-ARNt dans la région régulatrice d'ARNm, citons le cas de l'opéron His (Ames et al., 1983). En reconnaissant la structure pseudo-ARNt de son messenger, la MetRS (Dardel et al., 1990) l'AspRS de *S. cerevisiae* (Frugier et Giegé, 2003) (**figure 2.b**) et la ThrRS d'*E. coli* (Graffe et al., 1992 ; Springer et al., 1986 ; Torres-Larios et al., 2002 ; Romby et Springer, 2003) régulent leur propre

expression. Classiquement, la région 5' non traduite du messager se replie pour former une tige-boucle mimant la branche anticodon de l'ARNt correspondant. En absence d'ARNt, l'aaRS se fixe sur cette structure, ce qui empêche la fixation du ribosome et inhibe par conséquent la traduction.

Dans certains génomes de virus de plantes, des structures se repliant sous la forme de pseudo-ARNt sont présentes à leur extrémité 3' (**figure 2.c**, revue par Fechter et al., 2001). Ces structures sont aminoacylables par diverses synthétases (HisRS, ValRS ou TyrRS) avec des constantes cinétiques proches de celles des ARNt classiques. Elles interviennent dans l'initiation de la réplication de l'ARN viral (Dreher, 1999) et, dans le cas du virus de la Mosaïque Jaune du Navet, cette structure est impliquée dans l'initiation interne de la traduction de la polyprotéine virale (Barends et al., 2003).

## 6. Evolution des ARNt



**Figure 3. Etapes évolutives de l'ARNt**

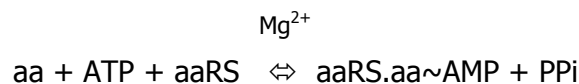
(a) Structure en tige boucle où les 4 nucléotides non appariés sont spécifiés : l'extrémité CCA universelle et la base discriminatrice D. ANT désigne l'anticodon. (b) Structure cruciforme issue de la duplication directe ou de la dimérisation de la tige boucle (Di Giulio, 1992). (c) Structure en double tige-boucle (Tanaka et Yo, 2001) correspondant à la structure secondaire de (b). (d) Structure secondaire classique de l'ARNt. ID désigne la région de l'ARNt portant les déterminants d'identité primordiaux et le triangle indique la position des introns dans les gènes d'ARNt. Figure adaptée de l'article de Di Giulio (2004).

L'ARNt est une molécule complexe dont la mise en place au cours de l'évolution a dû passer par plusieurs étapes. L'hypothèse actuellement admise est qu'elle est apparue par duplication d'une tige-boucle simple (Di Giulio, 1992 ; Schimmel et al., 1993) avant la mise en place des différentes spécificités et la subdivision des règnes du vivant (Widmann et al., 2005). La duplication d'une telle tige boucle, structure communément présente dans le monde ARN, permet d'aboutir à l'ARNt complet en passant par plusieurs étapes (**figure 3.b et 3.c**). La redistribution au cours de l'évolution des paires de bases impliquées dans la

formation des deux tiges boucles (Tanaka et Yo, 2001) (**figure 3.c**) aurait amené à la structure classique de l'ARNt (**figure 3.d**). Ce modèle d'évolution de l'ARNt implique des transitions évolutives de la synthèse protéique. Les précurseurs en tige-boucles devaient présenter un anticodon à leur extrémité 3' (**figure 3.a** : ANT) impliqué dans l'aminoacylation (Di Giulio, 1999) et dont les éléments d'identité primordiaux (**figure 3.d** : ID) constitueraient la trace (Möller et Janssen, 1990). Deux de ces structures, l'une flanquée de la chaîne polypeptidique en cours de synthèse, l'autre chargée avec un acide aminé devaient interagir avec l'ARNm de manière analogue au processus actuel pour permettre l'élongation de la chaîne polypeptidique. De la même manière, la structure en double tige-boucle aurait pu intervenir dans la synthèse protéique. Finalement, l'avènement de l'ARNt sous sa forme définitive (**figure 3.d**) aurait amené la synthèse protéique à sa forme actuelle avec d'éventuels intermédiaires. Ce modèle qui place un anticodon, donc une séquence spécifique, à proximité du site de fixation de l'acide aminé est en accord avec la présence des éléments d'identité primordiaux dans la branche acceptrice de l'ARNt (**cf. C.2**). Le mécanisme d'assemblage des deux structures en tige-boucle peut s'expliquer par l'association de deux mini-gènes par un mécanisme d' "exon shuffling" (Gilbert et al., 1997). Ceci est confirmé par la localisation des introns après l'anticodon dans les gènes d'ARNt, par l'existence d'ARNt codés par deux mini-gènes physiquement éloignés dans le génome de *Nanoaquifex equitans* (**cf. A.2.**) et par le repliement de l'ARNt sous la forme de deux hélices.

## B. Les aminoacyl-ARNt synthétases

Les aminoacyl-ARNt synthétases (aaRS) sont les enzymes qui catalysent l'aminoylation d'un acide aminé sur son ARNt homologue par une réaction en deux étapes :



La première étape, dite d'activation, décrite en 1955 par Hoagland (Hoagland, 1955) consiste à former un intermédiaire réactionnel avec l'acide aminé et une molécule d'ATP, intermédiaire qui sera ensuite transféré sur l'ARNt au cours d'une deuxième étape (Hoagland et al., 1957 ; Hoagland et al., 1958). Le transfert de l'acide aminé activé se fait sur le 2' ou 3' OH de l'adénosine terminale de l'ARNt. Certaines aaRS nécessitent l'ARNt pour l'étape d'activation de l'acide aminé, il s'agit de l'ArgRS, la GluRS, la GlnRS et la LysRS de classe I.

### 1. Les deux classes de synthétases

Les aaRS sont des enzymes modulaires de taille et d'oligométrie variées. La présence de motifs conservés au sein de leurs domaines catalytiques ainsi que leurs structures, mais aussi leurs propriétés fonctionnelles (face de l'ARNt reconnue et site de fixation de l'acide aminé) ont permis de les répartir en deux classes (Cusack et al., 1990 ; Eriani et al., 1990 ; Arnez et Cavarelli, 1997). Sous sa forme originale le tableau ne comportait pas les LysRS1, PylRS (pyrrolysyl-ARNt synthétase) et SepRS (*O*-phosphoséryl-ARNt synthétase) qui ont été ajoutées ultérieurement.

La LysRS est la seule aaRS à déroger à la règle des deux classes. Une LysRS de classe II (LysRS2) est présente chez la majorité des bactéries et eucaryotes, alors que les archaea et quelques bactéries possèdent une LysRS "spéciale" de classe I (LysRS1) (Ibba et al., 1997 ; Ambrogelly et al., 2002).

	<b>Classe I</b>	<b>Classe II</b>
<b>Caractéristiques</b>		
Motifs conservés	HIGH KMSKS	Motif 1 : ...P... Motif 2 : ...(Gx) <sub>3</sub> ER... Motif 3 : ...FRxE...
Site actif	feuillet de brins $\beta$ parallèles "Rossmann fold"	feuillet de brins $\beta$ antiparallèles
Conformation de l'ATP	étendu	replié
Liaison à l'ARNt	sillon mineur	sillon majeur
Conformation du CCA-3'	replié	étendu
Site d'aminocyclation	2' OH	3' OH
<b>Sous-classes</b>		
	<b>Ia</b>	<b>IIa</b>
	ValRS $\alpha$	SerRS $\alpha_2$
	IleRS $\alpha$	ThrRS $\alpha_2$
	LeuRS $\alpha$	ProRS $\alpha_2$
	MetRS* $\alpha_2$	GlyRS* $\alpha_2\beta_2$
	CysRS $\alpha$	HisRS $\alpha_2$
	ArgRS $\alpha$	AlaRS* $\alpha_4$
	<b>Ib</b>	<b>IIb</b>
	GluRS $\alpha$	AspRS $\alpha_2$
	GlnRS $\alpha$	AsnRS $\alpha_2$
	LysRS1 $\alpha$	LysRS2 $\alpha_2$
	<b>Ic</b>	<b>IIc</b>
	TrpRS $\alpha_2$	PheRS* $\alpha_2\beta_2$
	TyrRS $\alpha_2$	PyIRS #
		SepRS # $\alpha_4$

**Tableau 1. Principales caractéristiques des aaRS et répartition en classes et sous-classes**

Les séquences consensus HIGH et KMSKS des synthétases de classe I ne sont pas strictement conservées. Pour avoir une idée des variations possibles, se rapporter aux alignements de séquences de TyrRS dans la partie **Résultats**. La structure oligomérique est donnée pour les aaRS d'*E. coli*.

\* Synthétases dont la structure oligomérique n'est pas conservée dans l'évolution.

# La PyIRS et la SepRS sont codées uniquement dans certains génomes d'archaea méthanogènes. Ces deux aaRS particulières seront définies ultérieurement.

### 1.1. Les synthétases de classe I

La classe I comprend principalement des aaRS monomériques (8 en comptant la LysRS1) et seulement 3 dimériques. Leur site catalytique comprend un domaine de Rossmann (Rossmann et al., 1974), et deux séquences conservées dont le consensus est HIGH (His-Ile-Gly-His) et KMSKS (Lys-Met-Ser-Lys-Ser). Le domaine de Rossmann a initialement été identifié comme un motif de liaison aux nucléotides. Il est composé chez les aaRS d'un feuillet de 5 brins  $\beta$  parallèles reliés par des hélices  $\alpha$  et permet de lier l'ATP. Ce domaine renferme les séquences consensus HIGH et KMSKS et peut être interrompu par des

insertions CP1 et CP2. Les deux résidus histidine du motif HIGH sont impliqués dans la reconnaissance de l'acide aminé et de l'ATP. Les résidus lysine du motif KMSKS sont impliqués dans l'interaction avec l'ATP et localisés à proximité de l'extrémité acceptrice de l'ARNt. Ce motif se situe sur une boucle mobile qui change de conformation en fonction des substrats auxquels l'aaRS est liée. La subdivision des aaRS de classe I en sous-classes d'après leurs homologues de séquence correspond de manière surprenante à une classification des acides aminés en fonction de leur nature chimique. Ainsi, la sous-classe Ia correspond aux acides aminés hydrophobes, la sous-classe Ib aux acides aminés à longues chaînes latérales et la sous-classe Ic aux acides aminés aromatiques. Parmi les aaRS de classe I, plusieurs de la sous-classe Ia disposent d'un mécanisme de correction qui permet d'hydrolyser l'aa-AMP ou l'aa-ARNt quand un acide aminé incorrect est incorporé (**cf. C.1.**). Enfin, les structures cristallographiques de synthétases de classe I en complexe avec leur ARNt homologue montrent que l'interaction se fait du côté du grand sillon de la branche acceptrice. Cette propriété n'est néanmoins pas respectée par les deux aaRS de la sous-classe Ic, la TyrRS et la TrpRS qui interagissent avec leurs ARNt homologues par le petit sillon de la branche acceptrice, une propriété caractéristique des synthétases de classe II. De plus, ces dernières sont homodimériques, un type d'oligomérisation retrouvé majoritairement chez les aaRS de classe II.

## 1.2. Les synthétases de classe II

Les synthétases de classe II présentent des oligomérisations plus variées que les aaRS de classe I, allant de l'homodimère à l'hétérotétramère (à l'exception de la PheRS mitochondriale monomérique). La séparation des aaRS en deux classes a été établie grâce à la résolution de la structure de la SerRS d'*E. coli* en 1990 (Cusack et al., 1990) puis celle de l'AspRS de *S. cerevisiae* (Ruff et al., 1991) qui ont révélé la présence d'un nouveau motif structural pour reconnaître l'ATP. Trois séquences dégénérées sont présentes dans le domaine catalytique qui se replie en un feuillet de 9 brins  $\beta$  antiparallèles reliés par des hélices  $\alpha$ . Le motif 1, en plus de son rôle dans la formation du site catalytique, participe à la dimérisation de l'enzyme. Les motifs 2 et 3 participent aussi à la formation du site actif et présentent chacun un résidu arginine dans un contexte structural particulier. Ces aaRS ont aussi été subdivisées en sous-classes d'après leurs homologues de séquences. Les aaRS de la sous-classe IIa partagent la même organisation en feuillet  $\beta$  de leur domaine C-terminal à l'exception de la SerRS. Les trois aaRS de la sous-classe IIb conservent la même organisation de leur domaine de liaison à l'anticodon de l'ARNt en tonneau  $\beta$ . Enfin la classe

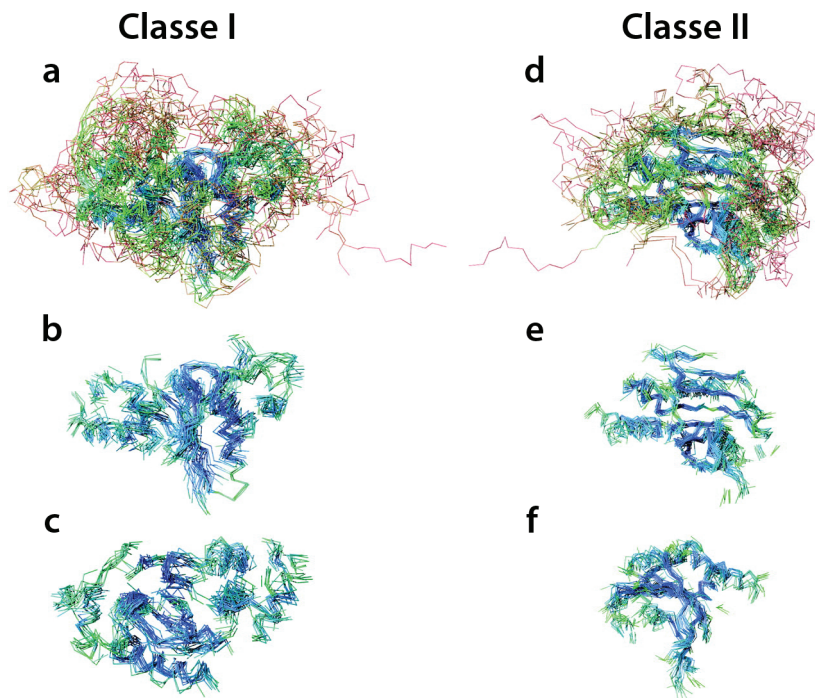


IIC est constituée par défaut des aaRS ne répondant pas aux critères des classes IIa et IIb. Néanmoins, la SepRS de *Methanococcus maripaludis* dont la structure cristallographique a récemment été résolue, montre que l'arrangement de ses sous-unités pour former l'homotétramère est très proche de celle de la PheRS hétérotétramérique (Kamtekar et al., 2007). Quant à la PylRS, son appartenance à la classe II a été décidée par analyse de sa séquence primaire, et son oligométrie devrait bientôt être révélée avec la résolution de sa structure (Yanagisawa et al., 2006). Rappelons enfin que les synthétases de la classe II ainsi que la TyrRS et la CysRS (de la classe I) fixent l'acide aminé sur le 3' OH du ribose de l'adénosine terminale de l'ARNt, à l'exception de la PheRS. Soulignons que le facteur EF-Tu qui prend en charge les ARNt aminoacylés pour les acheminer jusqu'au ribosome agit comme une isomérase et forme un orthoester avec des liaisons covalentes aux deux groupements hydroxyls du ribose (Förster et al., 1994).

## **2. Structure et organisation modulaire des aminoacyl-ARNt synthétases**

### **2.1. Organisation modulaire des aminoacyl-ARNt synthétases**

Les aaRS sont globalement composées d'un domaine catalytique "ancestral", responsable de l'activation et du transfert de l'acide aminé, et d'un domaine de liaison à l'anticodon de l'ARNt qui augmente la spécificité d'interaction entre les deux partenaires. Les nombreuses séquences et structures cristallographiques d'aaRS disponibles ont permis de caractériser les différents types de modules présents chez les synthétases. Aux deux domaines fonctionnels principaux viennent s'ajouter une soixantaine de modules additionnels impliqués soit dans la fonction d'aminocyclation de la synthétase soit dans l'édition (**cf. C.1.1.**). La présence de ces modules peut également conférer une nouvelle propriété à l'enzyme, indépendante de la synthèse protéique (**cf. B.4.**) (Wolf et al., 1999 ; O'Donoghue et Luthey-Schulten, 2003).

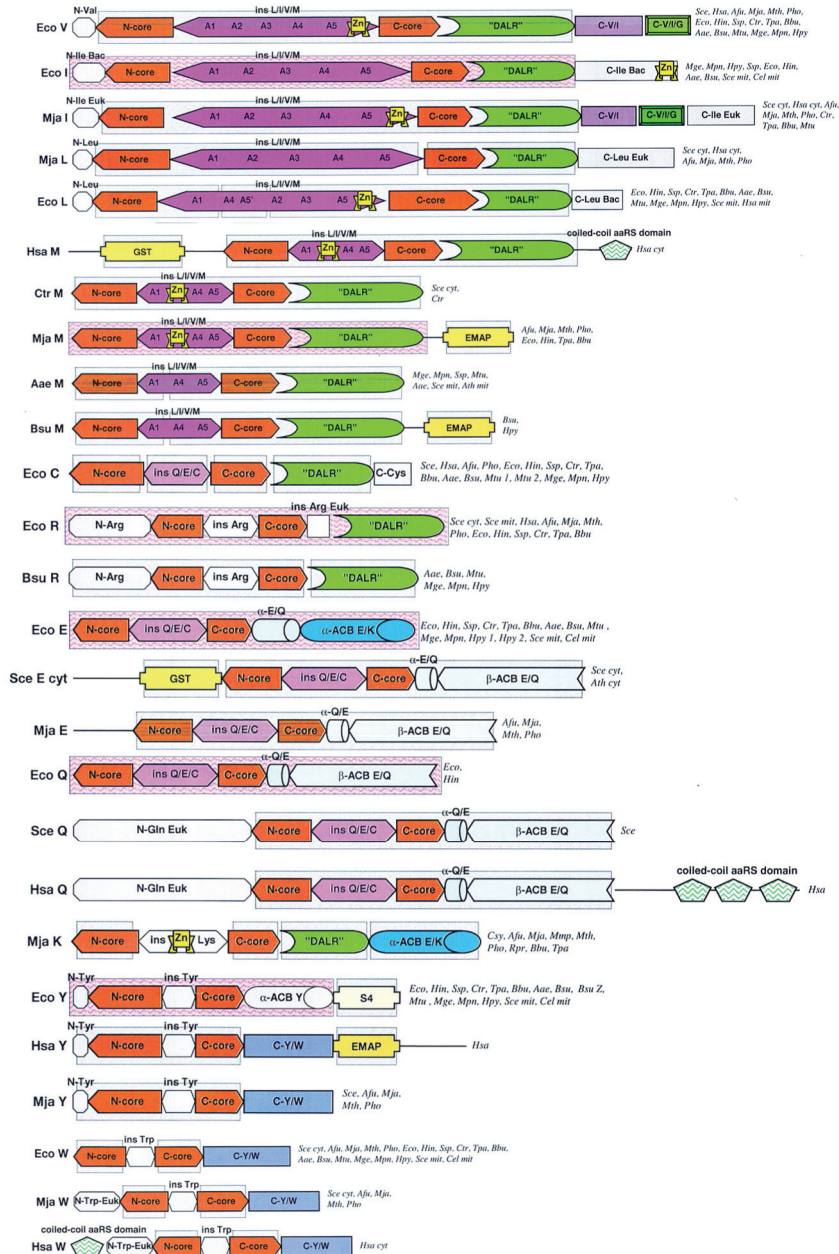


**Figure 4. Superposition des domaines catalytiques des aaRS de classes I et II**

Superposition des jeux de structures non redondants des domaines catalytiques des synthétases de classe I (**a** à **c**) et de classe II (**d** à **g**). Deux vues différentes des cœurs catalytiques conservés sont présentées dans les panneaux inférieurs. Les protéines sont colorisées d'après la conservation structurale. Les cœurs catalytiques pourraient refléter la structure des premières synthétases de classe I et II. Figure de l'article de O'Donoghue et Luthey-Schulten (2003).

Le module catalytique, tout d'abord, a une architecture très conservée mais différente entre les deux classes de synthétases, comme le montre la superposition des domaines catalytiques des aaRS de chaque classe (**figure 4**). Les domaines de liaison à l'anticodon sont de 7 types : 3 chez les aaRS de classe I, 3 chez les aaRS de classe II et 1 commun aux deux classes. Ceux des aaRS de classe I sont composés exclusivement d'hélices  $\alpha$  ou de feuillets  $\beta$  ( $\alpha$ -ACB E/K,  $\alpha$ -ACB Y,  $\beta$ -ACB E/Q), alors que ceux des aaRS de classe II sont composés d'une alternance d'hélices  $\alpha$  et de brins  $\beta$  ( $\alpha/\beta$ -ACB H/G/T/P, OB-ACB D/N/K, FDX-ACB). Le domaine de liaison à l'anticodon DALR (nommé d'après un motif caractéristique d'acides aminés conservés), exclusivement composé d'hélices  $\alpha$ , est présent dans 7 synthétases de la classe I (Arg, Cys, Lys, Met, Val, Ile, Leu) et la sous-unité  $\beta$  de la GlyRS (classe II).

la



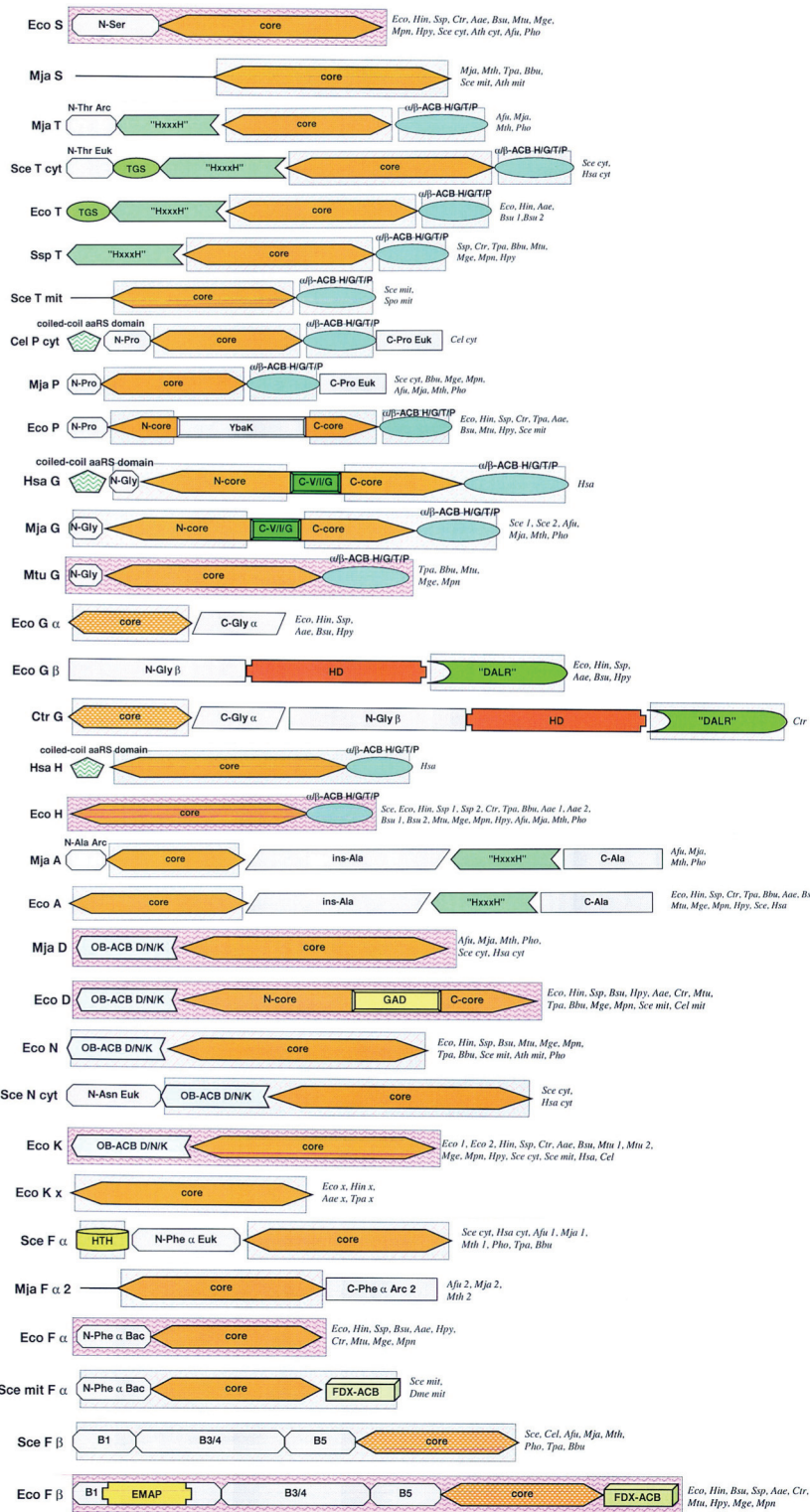
lb

lc

**Figure 5. Architecture modulaire des aminoacyl-ARNT synthétases de classe I**

L'architecture connue des différents domaines est indiquée pour chaque spécificité d'aaRS. Les noms abrégés des espèces dans lesquelles l'architecture modulaire dessinée a été observée sont précisés en bout de ligne. Les abréviations de noms de domaines sont : (A1-5) modules séparés identifiables dans les grandes insertions typiques des aaRS d'acides aminés aliphatiques ; (ACB) domaine de liaison à l'anticodon ; (FDX-ACB) ACB à repliement de type ferrédoxine ; (GST) glutathion *S*-transférase; (ins) insertion; (OB-ACB) ACB en "OB fold" ; (Zn) motif en doigt à zinc ; les autres domaines sont dénommés d'après la présence de séquences signatures conservées (DALR et HxxxH) ou d'après la spécificité des synthétases dans lesquels ils sont trouvés (C-V/I, C-V/I/G). Le fond rose indique les domaines pour lesquels la structure tridimensionnelle est déterminée; le fond bleu indique les domaines pour lesquels la structure peut être prédite d'après les similarités de séquences avec des domaines dont la structure est connue. (Aae) *A. aeolicus* ; (Afu) *A. fulgidus* ; (Ath) *A. thaliana* ; (Bbu) *B. burgdorferi* ; (Bsu) *B. subtilis* ; (Cel) *C. elegans* ; (Csy) *C. symbiosum* ; (Dme) *D. melanogaster* ; (Hsa) *H. sapiens* ; (Ctr) *C. trachomatis* ; (Eco) *E. coli* ; (Hin) *H. influenzae* ; (Hpy) *H. pylori* ; (Mja) *M. jannaschii* ; (Mge) *M. genitalium* ; (Mpn) *M. pneumoniae* ; (Mth) *M. thermoautotrophicum* ; (Mtu) *M. tuberculosis* ; (Mpn) *M. pneumoniae* ; (Pho) *P. horikoshii* ; (Rpr) *R. prowazekii* ; (Sce) *S. cerevisiae* (Spo) *S. pombe* ; (Ssp) *Synechocystis* sp. (Tpa) *T. pallidum*. Figure de l'article de Wolf et al. (1999).

IIa



IIb

IIc

**Figure 6. Architecture modulaire des aminoacyl-ARNt synthétases de classe II**  
 Pour les annotations, se référer à la légende de la figure 5.

Les autres modules fréquemment présents chez les synthétases peuvent être subdivisés en domaines soit spécifiques d'une classe d'aaRS, soit communs aux deux. Pour les domaines communs, on trouve (i) le domaine EMAP (Endothelial Monocyte Activating

Peptide) qui lie l'ARN avec sa structure dérivée du repliement en "OB fold" (Oligonucleotide/Oligosaccharide Binding fold), (ii) un domaine en hélices  $\alpha$  "coiled-coil" impliqué dans l'interaction avec l'ARNt et la formation du complexe multisynthétasique et (iii) un petit domaine C-terminal C-V/I/G. Les synthétases de classes I possèdent deux types d'insertions spécifiques ins Q/E/C et ins L/V/I/M aussi appelées domaine CP1, impliquées dans le second cas dans l'activité d'édition de la synthétase.

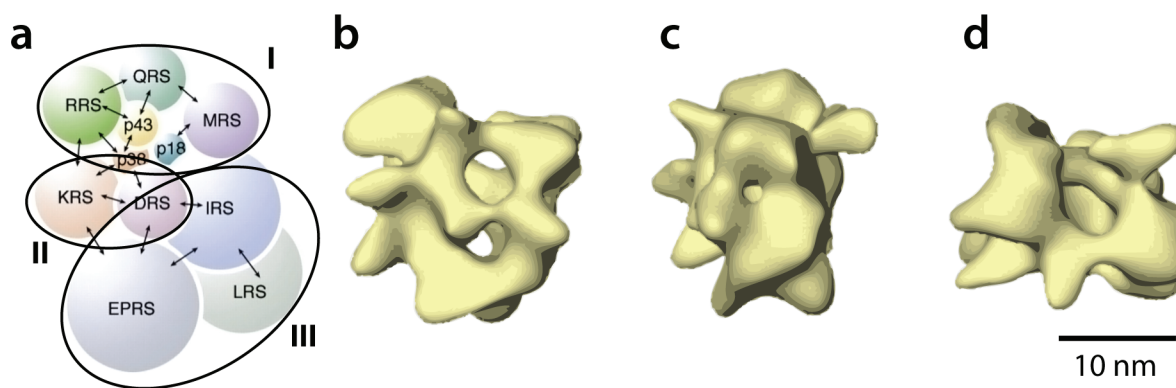
Différents domaines acquis par les synthétases au cours de l'évolution sont retrouvés dans d'autres types de protéines (Schimmel et Ribas de Pouplana, 2000). L'homologie entre ces domaines est souvent mieux conservée au niveau structural qu'au niveau de la séquence primaire. Citons le cas du domaine GAD inséré dans le cœur catalytique des AspRS eubactériennes qui est présent dans la ferrédoxine, la rubisco (ribulose-1,5-diphosphate) (Delarue et al., 1994) et la sous-unité GatE des Glu-ARNt<sup>Gln</sup> amidotransférases d'archaea (GatDE). Par ailleurs, le domaine YqeY à l'extrémité C-terminale de la GlnRS de *Deinococcus radiodurans* est présent à l'état libre dans de nombreux organismes ainsi que dans la sous-unité GatB des Glu-ARNt<sup>Gln</sup> amidotransférases eubactériennes (GatCAB).

Il est intéressant de noter qu'un certain nombre de domaines catalytiques des aaRS sont retrouvés dans des enzymes des voies de biosynthèse des acides aminés. Les protéines HisZ, par exemple, dérivent du domaine catalytique des HisRS. Elles ont perdu leur capacité d'aminocyclation, mais jouent un rôle essentiel dans la biosynthèse d'histidine (Sissler et al., 1999). De même, l'AsnA dérive du domaine catalytique de l'AspRS de levure et catalyse la transformation d'acide aspartique en asparagine (Nakatsu et al., 1998). Un autre exemple surprenant concerne le cas de la Glu-Q-RS d'*E. coli* qui est un paralogue du domaine catalytique de la GluRS d'*E. coli* et qui active puis transfère la Glu sur la queuosine en position 34 de l'ARNt<sup>Asp</sup> (Dubois et al., 2004).

## 2.2. Les complexes multisynthétasiques

La formation de complexes mettant en jeu plusieurs synthétases a été observée chez les archaea et les eucaryotes. Ainsi, chez l'archaea *Methanothermobacter thermoautotrophicus* un complexe ternaire entre la LysRS, la ProRS et la LeuRS améliore l'aminocyclation de l'ARNt<sup>Pro</sup> et de l'ARNt<sup>Lys</sup> par leurs aaRS respectives (Praetorius-Ibba et al., 2005; Praetorius-Ibba et al., 2007). De la même façon le complexe MetRS, GluRS et Arc1p présent chez *S. cerevisiae* stimule l'activité d'aminocyclation des deux synthétases (Simos et al., 1998).





**Figure 7. Le complexe multisynthétasique de mammifères**

(a) Schéma d'un modèle hypothétique d'organisation interne du complexe (Lee, 2004). (b-d) Structure tridimensionnelle du complexe obtenue par électro-microscopie (Wolfe et al., 2003). (b) Vue de face, (c) vue de profil par rotation de  $-90^\circ$  autour de l'axe vertical et (d) vue du dessus par rotation de  $-90^\circ$  autour de l'axe horizontal.

Chez les mammifères 9 des 20 synthétases (AspRS, LysRS, ArgRS, GlnRS, MetRS, LeuRS, IleRS et GluProRS) sont impliquées avec 3 protéines multifonctionnelles (p43, p38 et p18 renommées AIMP 1, 2 et 3) dans la formation d'un complexe supramoléculaire (Park et al., 2005a). Les AIMP (ARS-Interacting Multi-Functional Protein) participent toutes les trois à la régulation cellulaire en tant que molécules "signal". Le complexe multisynthétasique est connu et étudié par différentes approches pour en découvrir le réseau d'interactions et l'organisation structurale (par exemple Kellermann et al., 1982 ; Wolfe et al., 2005). Un modèle d'organisation proposé fait état de trois sous-domaines (**figure 7.a**). L'AspRS, la MetRS et la GlnRS constitueraient le premier sous-domaine, le second comprendrait la LysRS et l'ArgRS et le troisième, la GluProRS, l'IleRS et la LeuRS. Dans ce dernier sous-domaine, l'IleRS serait positionnée au centre et relierait les deux autres enzymes. Les trois AIMP connecteraient les sous-domaines entre eux et stabiliseraient l'ensemble de la structure. AIMP1 se situe au milieu du complexe et s'associe à l'ArgRS *via* sa région N-terminale. AIMP2 se lie à de nombreux composants et son rôle est critique pour l'assemblage du complexe (Kim et al., 2002). Les AIMP sont mutuellement dépendantes pour leur stabilité cellulaire (Han et al., 2006). Chaque composant semble lié au complexe différemment, probablement pour pouvoir répondre spécifiquement aux signaux déclenchant sa dissociation. Le rôle de ce complexe est encore mal compris, mais deux hypothèses peuvent être proposées. Tout d'abord, il pourrait améliorer l'efficacité et le contrôle de la traduction en canalisant le flux d'ARNt à aminoacyler dans des compartiments cellulaires ciblés. Ensuite, il pourrait servir de réservoir moléculaire pour le contrôle des activités de signalisation des synthétases qui le composent (**cf. B.4.4.**).

### 2.3. Les aminoacyl-ARNt synthétases dupliquées

Les cas de duplication d'aaRS sont rares et se trouvent majoritairement chez les bactéries. Ce phénomène peut avoir deux finalités différentes : il peut permettre à l'organisme de palier l'absence d'une autre synthétase, c'est le cas de l'AspRS et de la GluRS qui sera décrit dans le paragraphe suivant, ou d'acquérir grâce à la deuxième version de l'aaRS une résistance accrue dans des conditions de stress. Ce second phénomène concerne les TyrRS, ThrRS, IleRS, MetRS, TrpRS et LysRS eubactériennes. Les streptomyces qui produisent l'indolmycine possèdent une seconde TrpRS inductible et résistante à l'antibiotique (Kitabatake et al., 2002). *D. radiodurans* possède aussi deux versions de TrpRS, l'une à spécificité stricte, l'autre à spécificité relâchée potentiellement impliquée dans une voie de signalisation (Buddha et al., 2004). Le cas le plus intrigant de duplication fonctionnelle implique la LysRS. Chez *E. coli*, il existe deux versions de la LysRS2 : une version *lysS* constitutive et une version *lysU* inductible par choc thermique. La seconde LysRS2 est résistante à l'inhibition par des analogues de la lysine tels que la cadavérine, ce qui permet une traduction efficace dans des conditions de stress (Brevet et al., 1995). Par ailleurs, on trouve chez des archaea et des bactéries pathogènes à la fois la LysRS2 et une LysRS structuralement différente de classe I (LysRS1), plus sélective et moins sensible à l'inhibition (Levengood et al., 2004).

## 3. Evolution des aminoacyl-ARNt synthétases

La répartition des synthétases en deux classes équivalentes de 10 membres semble trop homogène pour être fortuite. L'analyse de leur séquence et la possibilité de fixer une synthétase de chaque classe de part et d'autre d'un même ARNt (Ribas de Pouplana et Schimmel, 2001) ont amené à proposer que les deux domaines catalytiques primordiaux des synthétases dérivent des brins sens et anti-sens d'un même gène (Rodin et Ohno, 1995; Rodin et Rodin, 2006; Pham et al., 2007). Une autre théorie propose une apparition indépendante des deux classes de synthétases qui seraient les traces de deux systèmes de traduction alternatifs (Nagel et Doolittle, 1991). Les synthétases de classe II sont supposées antérieures à celles de classe I car (i) elles aminoacylent des minihélices, (ii) leur ARNt homologues présentent des traces de l'anticodon dans la branche acceptrice (Möller et Janssen, 1990), et (iii) elles sont spécifiques d'acides aminés simples et petits, dont la plupart jouent un rôle structural important. Ce sont ces mêmes acides aminés qui étaient

générés dans l'expérience de Miller qui visait à recréer l'environnement terrestre primordial (Trifonov et Berezovsky, 2002). Enfin, (iv) les codons riches en GC, supposés prépondérants dans les gènes ancestraux, sont plus fréquents dans les triplets codant les acides aminés des synthétases de classe II (Ferreira et Cavalcanti, 1997). Une étude récente a renforcé cette hypothèse en faisant le lien entre le domaine catalytique des aaRS de classe II et des enzymes des voies de biosynthèse d'acides aminés (Klipcan et Safro, 2004). Il y aurait ainsi eu une évolution concomitante entre le code génétique, les voies de biosynthèse des acides aminés et les synthétases.

Quoiqu'il en soit, toutes les études s'accordent sur la présence du jeu complet de synthétases de classe I et II ayant déjà acquis dans LUCA (Last Universal Common Ancestor, Kyrpides et al., 1999) les architectures modulaires proches de celles des synthétases actuelles (Wong, 2005). Par conséquent, l'évolution des deux classes de synthétases depuis leurs ancêtres communs incluant plusieurs cycles de duplication est antérieure à LUCA. Leur schéma évolutif global commun est en accord avec la phylogénie déterminée par l'analyse de l'ARNr 16S, mais chaque arbre phylogénétique d'aaRS est unique et révèle de nombreux événements de duplication et de transfert de gènes (Diaz-Lazcoz et al., 1998; Wolf et al., 1999; Woese et al., 2000; O'Donoghue et Luthey-Schulten, 2003; Brindefalk et al., 2007).

## **4. Fonctions non-canoniques des aminoacyl-ARNt synthétases**

### **4.1. Autorégulation de l'expression**

Grâce à leur capacité à interagir avec l'ARN, les aaRS sont capables d'autoréguler leur expression au niveau transcriptionnel ou traductionnel. Ainsi, l'AlaRS d'*E. coli* inhibe la transcription de son propre gène en se liant à une séquence palindromique dans le promoteur à proximité du site d'initiation de la transcription (Putney et Schimmel, 1981). La ThrRS d'*E. coli* se lie à la région 5' non traduite de son propre ARNm, ce qui inhibe l'initiation de la traduction (Springer et al., 1986). Dans la levure *S. cerevisiae* l'expression de l'AspRS est régulée par un processus mettant en jeu la reconnaissance de son propre ARNm et la sensibilité à la concentration en ARNt<sup>Asp</sup> (Frugier et al., 2005).



## 4.2. Epissage d'introns

Les TyrRS mitochondriales des levures *Neurospora crassa* et *Podospira anserina* interviennent dans l'épissage d'introns de groupe I (Akins et Lambowitz, 1987 ; Kamper et al., 1992). Ces derniers se replient sous une structure d'ARNt (Caprara et al., 1996) et, tout comme l'ARNt<sup>Tyr</sup>, se fixent à cheval sur les deux sous-unités de la synthétase homodimérique. Ils interagissent avec la TyrRS à l'opposé de la face qui reconnaît l'ARNt<sup>Tyr</sup> (Paukstelis et al., 2005). Le site de fixation des introns est créé par trois insertions spécifiques accompagnées de changements concertés des résidus voisins. Par ailleurs, la LeuRS mitochondriale de *S. cerevisiae* se lie à l'intron bI4 et forme un complexe avec la maturase bI4 pour cliver l'intron de groupe I (Labouesse, 1990).

## 4.3. Implication dans le cycle viral

L'ARNt<sup>Lys3</sup> qui sert d'amorce pour la rétro-transcription du génome du virus HIV de type 1 est incorporé dans les virions avec sa synthétase homologue, la LysRS (Cen et al., 2001). Cette incorporation de la synthétase est indépendante de l'ARNt et se fait par l'interaction avec la protéine virale GAG. Ce complexe binaire facilite ensuite l'incorporation de l'ARNt<sup>Lys</sup> grâce à l'interaction non productive entre l'ARNt et la LysRS (Cen et al., 2004).

## 4.4. Molécules "signal"

Un grand nombre de synthétases humaines sont impliquées dans des voies de signalisation. Ainsi, la LysRS est sécrétée en réponse à la cytokine pro-inflammatoire TNF- $\alpha$  et induit la production de TNF- $\alpha$  par les macrophages et leur migration (Park et al., 2005b). La GluRS inhibe une voie de signalisation pro-apoptotique (Ko et al., 2001) par son interaction avec ASK1 (kinase régulant le signal apoptotique). La TyrRS cytoplasmique humaine est sécrétée hors de la cellule et clivée par l'énolase en un fragment N-terminal pro-angiogénique, la mini-TyrRS, et un fragment C-terminal qui stimule la production de TNF- $\alpha$  par les macrophages et leur migration (Wakasugi et Schimmel, 1999). L'activité de la mini-TyrRS est similaire à celle des chémokines comme l'interleukine 8 et dépend de la présence d'un motif de trois acides aminés ELR dans le domaine catalytique (Wakasugi et al., 2002). À l'opposé, la TrpRS acquiert une activité anti-angiogénique par épissage alternatif (mini-TrpRS) ou par protéolyse (T2-TrpRS) (Otani et al., 2002). Cette activité inhibe la mini-TyrRS, la croissance de vaisseaux et les voies de signalisation de cellules

endothéliales activées suite à une lésion (Tzima et al., 2003). L'activité anti-angiogénique de la T2-TrpRS pourrait d'ailleurs être mise à profit pour bloquer la néo-vascularisation des tumeurs sans affecter les vaisseaux sains en complément d'une thérapie de cytotoxicité (Tzima et Schimmel, 2005).

#### **4.5. Voies métaboliques et régulation**

Plusieurs aaRS sont impliquées dans des voies métaboliques propres à l'organisme et des processus de régulation. Chez *S. cerevisiae*, la SerRS interagit avec la protéine Pex21p et intervient dans la biogenèse du péroxysome (Rocak et al., 2002; Godinic et al., 2007). La TyrRS, quant à elle, interagit avec Knr4p pour l'assemblage de la paroi externe (Dagkessamanskaia et al., 2001). Chez *Methanocaldococcus jannaschii*, la protéine Mj1338 s'associe à plusieurs aaRS dont la ProRS. Le rôle supposé de cette protéine dans le métabolisme monocarboné de l'archaea suggère l'existence d'un lien entre métabolisme et traduction (Lipman et al., 2003). Chez l'homme, la GluProRS (EPRS) est libérée du complexe multisynthétasique (**cf. B.2.2.**) par phosphorylation. Elle rejoint alors le complexe GAIT (Inhibiteur de la Transcription Activé par l'interféron Gamma, composé de la sous-unité ribosomale L13a et du GADPH). Sous cette forme, elle reconnaît et lie une structure secondaire de la région 3' non traduite des transcrits du facteur de croissance endothélial vasculaire et de la céruloplasmine (Sampath et al., 2004) et réprime leur traduction. Ce processus permet de limiter les inflammations et lésions tissulaires provoquées par l'accumulation de céruloplasmine.

#### **4.6. Implications dans des pathologies**

Certaines synthétases sont impliquées dans des pathologies soit suite à des mutations de leur gène soit en tant que cible dont l'activité enzymatique est inhibée par la pathologie.

A ce jour, plusieurs synthétases humaines ont été identifiées comme directement impliquées dans des pathologies. Il a été montré qu'une mutation "non-sens" dans le domaine d'édition de l'AlaRS cytoplasmique est associée à une ataxie et à une perte de cellules de Purkinje (Lee et al., 2006). L'accumulation de protéines mal repliées entraîne la mort des cellules nerveuses, ce qui rend cette pathologie neurodégénérative. Par ailleurs, des mutations du gène de l'AspRS mitochondriale affectant son activité causent une leucoencéphalopathie impliquant le tronc cérébral et la moelle épinière (Scheper et al.,

2007). Dans le cas de la GlyRS cytoplasmique, une mutation de son gène a été reliée à la neuropathie de Charcot Marie Tooth (CMT) de type 2D, une maladie héréditaire qui affecte le système nerveux périphérique (Antonellis et al., 2003). Les malades présentent des jonctions neuromusculaires de morphologie anormale, une transmission du signal neuronal perturbée et moins rapide, ainsi qu'une réduction des axones périphériques. Des analyses fonctionnelles ont montré que cette mutation n'affecte pas la fonction d'aminacylation, mais la localisation de la synthétase dans les cellules nerveuses. De ce fait, des déficits locaux de synthèse protéique ont été détectés (Antonellis et al., 2006 ; Seburn et al., 2006). De même, une mutation dans le gène de la TyrRS cytoplasmique est impliquée dans la CMT intermédiaire progressive de type C (Jordanova et al., 2006). Ici, la mutation affecte non seulement la localisation de la synthétase, mais aussi sa fonction d'aminacylation, ce qui aboutit à l'incapacité de répondre à la forte demande de synthèse protéique du système nerveux périphérique. Ceci pourrait avoir un effet sur la plasticité des synapses aboutissant à la dégénérescence puis à la perte des axones et finalement à une neuropathie périphérique motrice et sensorielle.

La première synthétase humaine identifiée comme cible d'auto-anticorps dans une myosite est l'HisRS (Mathews et Bernstein, 1983). Cette maladie auto-immune se caractérise par une faiblesse et une perte musculaire. Depuis, 7 autres synthétases ont été identifiées comme cibles d'auto-anticorps : la ThrRS (Mathews et al., 1984), l'AlaRS (Bunn et al., 1986), l'IleRS (Targoff, 1990), la GlyRS (Targoff, 1990), l'AsnRS (Hirakata et al., 1999), la TyrRS (Hashish et al., 2005) et la PheRS (Betteridge et al., 2007). Les manifestations cliniques sont variées et ne permettent pas de définir clairement un type de pathologie associée. Par ailleurs, on sait aussi que chez des patients atteints de sclérose latérale amyotrophique, la superoxyde dismutase mutée se lie à la LysRS (Kunst et al., 1997). Enfin, l'AsnRS du parasite *Brugia malayi* responsable de la filariose lymphatique chez l'Homme a des propriétés chimiotactiques pour les neutrophiles et les éosinophiles, ce qui contribue au développement de réactions inflammatoires chroniques (Ramirez et al., 2006).

#### **4.7. Les aminoacyl-ARNt synthétases comme cibles thérapeutiques**

La découverte des propriétés antibiotiques de la pénicilline par Fleming en 1928 a ouvert la voie à la lutte contre de nombreuses maladies jusqu'alors incurables. De nos jours, les antibiotiques sont d'usage courant, mais leur surconsommation a entraîné l'apparition de résistances chez de nombreux agents pathogènes, et même des multirésistances (citons les exemples de la bactérie *Staphylococcus aureus* responsable d'infections nosocomiales ou

encore le parasite *Plasmodium falciparum* causant le paludisme). Ce constat souligne l'importance de la mise au point de nouvelles générations d'antibiotiques. Malgré leur rôle central dans le processus de traduction, les aaRS sont encore très peu exploitées comme cibles thérapeutiques. Elles présentent pourtant plusieurs avantages : (i) elles ont acquis une grande diversité structurale qui est exploitable dans la mise au point d'inhibiteurs spécifiques d'une espèce, (ii) le choix de la cible pour bloquer l'aminacylation est large parmi la vingtaine de synthétases disponibles, (iii) les données structurales et fonctionnelles sont de plus en plus nombreuses et permettent la mise au point rationnelle d'inhibiteurs, (iv) ces inhibiteurs auraient une faible réactivité vis-à-vis des antibiotiques déjà existants limitant les risques d'interactions médicamenteuses, et (v) il existe déjà un certain nombre d'inhibiteurs naturels pouvant servir de point de départ pour la mise au point d'antibiotiques efficaces, spécifiques et utilisables cliniquement (Kim et al., 2003; Hurdle et al., 2005; Ochsner et al., 2007). Une dizaine de ces inhibiteurs naturels sont connus : l'indolmycine et la chuangxinmycine contre la TrpRS, la borrelidine contre la ThrRS, la granaticine contre la LeuRS, la furanomycine contre l'IleRS, l'ochratoxine A contre la PheRS, la cispentacine contre la ProRS et la mupirocine contre l'IleRS. Ce dernier inhibiteur a donné lieu à de nombreuses études et est le seul actuellement exploité commercialement (Bactroban, GlaxoSmithKline). Pour éviter le développement de souches bactériennes résistantes à ces antibiotiques la recherche d'inhibiteurs ciblant plusieurs synthétases simultanément est une voie à explorer.

## C. Spécificité de la réaction d'aminacylation

La fidélité de la traduction repose sur deux associations successives : celle de l'acide aminé et de l'ARNt et celle de l'interaction ARNt/ARNm au niveau du ribosome. L'aaRS associe un acide aminé à un ARNt et doit être spécifique pour ces deux types de substrats à la fois. La sélection de l'acide aminé est finement contrôlée et des mécanismes ont été développés pour rejeter les acides aminés incorrectement activés. La sélection de l'ARNt se fait par la reconnaissance de signaux sur l'ARNt qui constitueraient un "second code génétique" (Lengyel, 1966; de Duve, 1988).

### 1. Spécificité pour l'acide aminé

La possibilité d'erreurs lors de la synthèse protéique due à l'incorporation d'un acide aminé incorrect a été postulée dès les années 50 (Pauling, 1957). Alors que le taux d'activation incorrecte théorique pour des acides aminés isostériques ne différant que par un seul groupe méthyl (comme la valine et l'isoleucine) serait de 1 pour 5, le taux mesuré *in vivo* n'est que de 1 pour 3000 (Loftfield, 1972). Ainsi, l'existence de mécanismes permettant d'améliorer la discrimination des acides aminés par les synthétases a été proposée. La nature du site actif des aaRS permet une première étape de sélection en rejetant les acides aminés plus encombrés. Cependant il ne discrimine pas les acides aminés isostériques et plus petits que son acide aminé spécifique. Pour palier ce problème, un site d'édition a été mis en place dans les aaRS qui hydrolysent les aa-AMP incorrects (Fersht, 1977).

#### 1.1. Édition des erreurs d'activation ou d'acylation d'acides aminés

L'incorporation d'acides aminés incorrects dans les protéines en cours de synthèse peut avoir des effets dramatiques sur la cellule. Il a récemment été mis en évidence qu'un défaut du mécanisme d'édition de l'AlaRS entraîne l'accumulation de protéines mal repliées et la mort de la cellule (Lee et al., 2006). Le rejet de l'acide aminé peut avoir lieu à deux niveaux : soit après formation de l'aminacyl-adenylate, on parle alors d'édition pré-transfert (Fersht, 1977), soit après formation de l'aminacyl-ARNt, on parle alors d'édition post-transfert (Eldred et Schimmel, 1972).

Le premier cas étudié est celui de l'IleRS qui édite le Val-AMP de façon ARNt<sup>Ile</sup> dépendante (Baldwin et Berg, 1966) et le Val-ARNt<sup>Ile</sup> (Eldred et Schimmel, 1972). Des activités similaires ont été observées pour deux autres aaRS de classe I, la LeuRS et la ValRS, qui se localisent dans un domaine spécifique, différent et éloigné du domaine catalytique, appelé domaine CP1 (Starzyk et al., 1987). Dans le cas de l'IleRS, il a été démontré par mutagenèse que le site CP1 présente deux sites de fixations différents pour l'édition de la valine, pré- ou post-transfert (Hendrickson et al., 2002). Les aaRS de classe II qui ont une activité d'édition (ThrRS, AlaRS, ProRS et PheRS) ne possèdent pas de domaine d'édition commun. Celui-ci est différent suivant les synthétases et les organismes. La ThrRS, par exemple, présente des domaines d'édition phylogénétiquement et structuralement différents entre eucaryotes/bactéries et archaea (Dock-Bregeon et al., 2000 ; Dock-Bregeon et al., 2004 ; Hussain et al., 2006). Néanmoins le mécanisme d'édition post-transfert des synthétases de classe II semble être conservé du point de vue fonctionnel (Ling et al., 2007). Les deux classes de synthétases mettent à profit la capacité de l'extrémité CCA acceptrice de l'ARNt à permuter entre une conformation hélicoïdale ou en coude pour remplir leur fonction d'aminocyclation et d'édition. Etant donné la fixation de l'ARNt par l'une ou l'autre de ses faces suivant la classe de synthétase, la permutation du CCA est aussi inversée entre les deux classes d'aaRS (Silvian et al., 1999 ; Dock-Bregeon et al., 2000 ; Fukai et al., 2000). Dans le cadre des premières étapes de l'évolution des systèmes vivants, l'incorporation d'acides aminés incorrects pouvait avoir un avantage évolutif pour des cellules poussant dans des milieux carencés en certains acides aminés, car elle permet de poursuivre la synthèse de la chaîne polypeptidique, mais aussi comme source de variabilité pour la synthèse de nouvelles protéines (Pezo et al., 2004).

## 1.2. Les domaines d'édition à l'état libre

Il existe dans de nombreux organismes des protéines homologues aux domaines d'édition des aaRS qui sont impliquées dans l'édition en *trans* d'ARNt chargés avec un acide aminé incorrect. Ainsi, la protéine YbaK de *Haemophilus influenzae* est capable d'hydrolyser spécifiquement le Cys-ARNt<sup>Pro</sup> en interagissant avec la ProRS (An et Musier-Forsyth, 2004). La protéine PrdX de *Clostridium sticklandii*, homologue du domaine d'édition des ProRS, hydrolyse spécifiquement et efficacement l'Ala-ARNt<sup>Pro</sup>. De la même manière, les protéines AlaX de *Methanosarcina barkeri* et *Sulfolobus solfataricus*, homologues du domaine d'édition de l'AlaRS, hydrolysent le Ser-ARNt<sup>Ala</sup> et le Gly-ARNt<sup>Ala</sup>, respectivement (Ahel et al., 2003 ; Fukunaga et Yokoyama, 2007a). Une D-Tyr-ARNt<sup>Tyr</sup> déacylase (DTD) a été identifiée chez

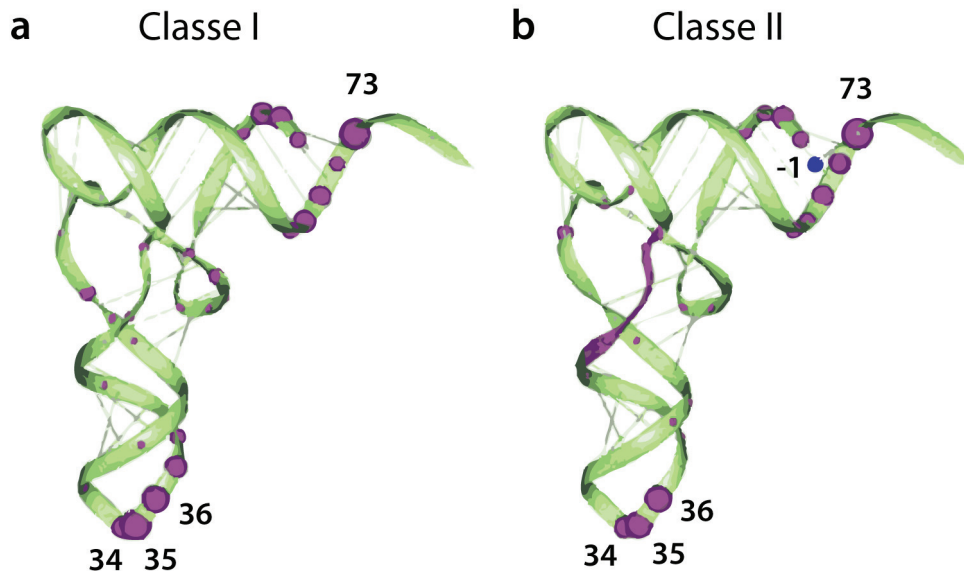
*E. coli* (Calendar et Berg, 1967) et hydrolyse le D-Tyr-ARNt<sup>Tyr</sup> mais aussi le D-Trp-ARNt<sup>Trp</sup>, D-Asp-ARNt<sup>Asp</sup>, D-Ser-ARNt<sup>Ser</sup> et D-Glu-ARNt<sup>Glu</sup> (Soutourina et al., 2000). Cette protéine est retrouvée chez les bactéries et les eucaryotes, mais est absente chez les archaea où l'activité déacylase est assurée par la protéine DTD2 (Ferri-Fioni et al., 2006).

## 2. Spécificité pour l'ARNt

La fidélité de la réaction d'aminacylation repose également sur la spécificité de reconnaissance entre ARNt et synthétase homologue (ou les ARNt de la même famille isoacceptrice). Cette seconde étape n'est pas liée uniquement à la reconnaissance du substrat, mais également à la cinétique de sa charge. En effet, l'affinité de l'aaRS pour son ARNt homologue n'est qu'une centaine de fois supérieure à celle pour les ARNt hétérologues (Ebel et al., 1973). La reconnaissance par l'aaRS d'éléments structuraux distribués sur l'ARNt induit des changements conformationnels qui aboutissent à une interaction productive. Les éléments de l'ARNt permettant la reconnaissance par l'aaRS sont appelés des "déterminants" alors que ceux qui défavorisent l'interaction avec une aaRS hétérologue sont appelés des "antidéterminants". L'ensemble de ces éléments constitue le "jeu d'identité" de l'ARNt.

Les jeux d'identités des ARNt ont été déterminés complètement chez *E. coli* et *S. cerevisiae* et ponctuellement chez un certain nombre d'organismes modèles, notamment l'eubactérie hyperthermophile *Thermus thermophilus*, le nématode *Caenorhabditis elegans*, la plante *Arabidopsis thaliana* et l'Homme (revue par Giegé et al., 1998). Les connaissances accumulées dans tous ces systèmes ont permis de dégager un certain nombre de règles générales (revues par McClain, 1995; Giegé et al., 1998; Beuning et Musier-Forsyth, 1999; Giegé et Frugier, 2003). (i) Les éléments d'identité se localisent préférentiellement aux extrémités (anticodon et acceptrice) de la structure en "L" de l'ARNt (**figure 8**). (ii) Un jeu d'identité est constitué de la combinaison des différents éléments d'identité dans un contexte structural et de séquence particulier. (iii) Les jeux d'identité sont généralement conservés au cours de l'évolution pour une même spécificité. (iv) Un élément d'identité peut être un nucléotide (par exemple la base discriminatrice A73 de l'ARNt<sup>Tyr</sup>; Himeno et al., 1990), une paire de bases (G3:U70 de l'ARNt<sup>Ala</sup>; Hou et Schimmel, 1988), un élément structural (l'orientation de la grande région variable de l'ARNt<sup>Ser</sup>; Himeno et al., 1990; Lenhard et al., 1999) ou une base modifiée (lysidine de l'ARNt<sup>Ile</sup>; Muramatsu et al., 1988). (v) Un jeu d'identité est complet quand sa transplantation dans un ARNt d'une autre spécificité confère une nouvelle spécificité pour l'ARNt hôte. Néanmoins, un affinement structural est parfois

nécessaire pour une présentation optimale des éléments de l'ARNt hôte. Ceci est possible notamment en faisant varier le nombre de nucléotides présents de part et d'autre des G18 et G19 conservés de la boucle D (Perret et al., 1992).



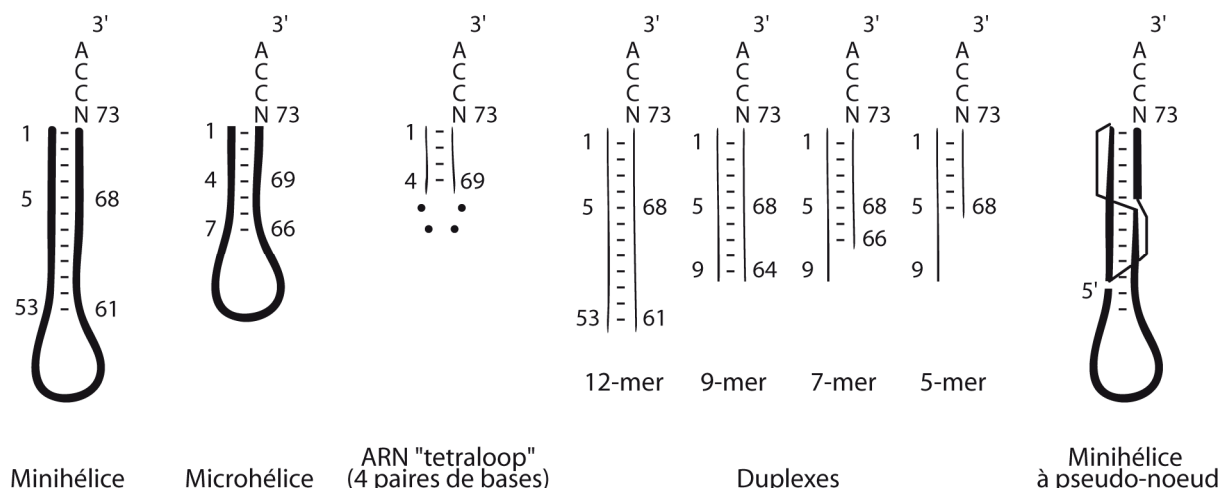
**Figure 8. Localisation et importance des nucléotides des jeux d'identité des ARNt d'*E. coli***  
 La taille des sphères est proportionnelle à la fréquence de nucléotides d'identités à une position donnée. La région variable colorée en violet indique sa participation dans l'identité Ser, Tyr et Leu. Les éléments sont placés sur la structure de l'ARNt<sup>Phe</sup> (extrait de Giegé et al., 1998).

Les jeux d'identités des ARNt ont été déterminés par des approches complémentaires *in vivo* (Normanly *et al.*, 1986; Gabriel et McClain, 1999) et *in vitro*. Dans les deux cas, la stratégie consiste à muter les positions d'intérêt de l'ARNt et à observer les effets de la mutation sur la réaction d'aminacylation. L'étude *in vitro* est basée sur la mesure des paramètres cinétiques d'aminacylation de molécules d'ARNt transcrites par la T7 ARN polymérase et, de ce fait, dépourvues de modifications post-transcriptionnelles (Sampson et Uhlenbeck, 1988). La comparaison de l'efficacité catalytique (rapport  $k_{cat}/K_M$ ) d'aminacylation des molécules sauvages et mutées permet de déterminer les éléments d'identité de l'ARNt étudié. Cette méthode est sensible et permet de tester de manière systématique les nucléotides de l'ARNt. Néanmoins, elle ne s'applique qu'aux ARNt fonctionnels sans modifications post-transcriptionnelles et ne permet pas de reproduire le contexte de compétition avec les autres constituants cellulaires. Certaines limitations sont dépassées par l'utilisation de techniques *in vivo* qui mesurent l'effet des mutations de l'ARNt sur son aminacylation dans le contexte cellulaire. Une des approches est basée sur la capacité d'un ARNt suppresseur à incorporer un acide aminé au niveau du codon stop d'une protéine. L'ARNt ainsi produit *in vivo* possèdera, outre des mutations ponctuelles propres à



l'ARNt naturel étudié, un anticodon capable de s'adapter au codon stop, d'où le nom d'ARNt suppresseur (Normanly et al., 1990). Enfin, on peut utiliser des souches dont on a supprimé le gène d'un ARNt (ou une famille isoacceptrice d'ARNt) et qui sont complétées par l'ARNt sauvage ou muté à étudier (Gabriel et McClain, 1999). Ces méthodes ne permettent de détecter que les nucléotides qui ont un rôle prépondérant dans l'identité de l'ARNt. De plus, la mutation de l'ARNt peut avoir un effet sur d'autres étapes cellulaires, rendant l'interprétation des résultats délicate. Rappelons enfin qu'un élément d'identité n'est pas obligatoirement en contact direct avec la synthétase et inversement les résidus en contact avec des acides aminés dans les structures cristallographiques de complexes ARNt/aaRS ne font pas toujours partie du jeu d'identité.

### 3. Aminoacylation de mini substrats



**Figure 9. Structures aminoacylables minimalistes dérivées du bras accepteur d'ARNt**

Figure adaptée de Schimmel et al., 1993.

La localisation confinée des nucléotides d'identité a permis la conception de molécules récapitulant différents domaines de l'ARNt et qui sont aminoacylables. Le premier cas étudié à été celui de minihélices de 12 paires de bases qui reproduisent la branche acceptrice de l'ARNt<sup>Ala</sup> de *E. coli* (Francklyn et Schimmel, 1989). Dans ce cas, c'est la présence de la paire de base G3-U70, élément d'identité majeur, qui permet l'aminoacylation efficace et spécifique de ces minihélices par l'AlaRS. Des minihélices mimant les branches acceptrices d'ARNt de plusieurs autres spécificités sont aussi aminoacylables par leur aaRS homologue. La réduction de ces minihélices en microhélices de seulement 7 paires de bases leur confère toujours la capacité d'être aminoacylées, mais une efficacité fortement réduite

(**Figure 9**, revue par Schimmel et al., 1993). Enfin, citons le cas le plus extrême de l'aspartylation d'une structure minimale composée de quatre paires de bases fusionnées à une extrémité GCCA-3' (Frugier et al., 1994). Cette aminoacylation est spécifique de la présence du nucléotide G73 qui est reconnu par l'AspRS de la même manière que dans le contexte de l'ARNt<sup>Asp</sup>. Ce type de molécules pourrait être les substrats originels des synthétases, qui auraient amélioré la discrimination entre les différentes spécificités par l'acquisition de la branche anticodon et du domaine de reconnaissance de l'anticodon.

#### 4. Systèmes d' aminoacylation artificiels

L'incorporation d'acides aminés non-naturels dans des protéines est un outil de choix pour étudier la structure, la localisation et les interactions entre protéines grâce à des groupements réactifs tels que des sondes spectroscopiques, des fluorophores ou des chélatants de métaux. Plusieurs méthodes ont été développées pour y parvenir. La première consiste à reprogrammer un codon stop (ou plus récemment un quadruplet de bases; Sisido et al., 2005) pour incorporer un acide aminé non naturel spécifique grâce à un couple ARNt/aaRS génétiquement modifié (revue par Wang et al., 2006). L'ARNt a été muté pour présenter l'anticodon correspondant au codon stop ciblé et est capable de fixer l'acide aminé sur son extrémité 3'. La synthétase reconnaît et active spécifiquement l'acide aminé non-naturel, l'ARNt muté et catalyse la fixation de l'acide aminé sur l'ARNt. Pour assurer la spécificité de la réaction et ne pas perturber la synthèse du reste de la chaîne, le couple ARNt/aaRS dit "orthogonal" doit être spécifique et ne pas réagir avec les autres systèmes d' aminoacylation : aucune synthétase endogène ne doit reconnaître cet ARNt "orthogonal" et aucun ARNt endogène ne doit être reconnu par la synthétase "orthogonale". Le système tyrosine est l'un des plus intéressants pour la mise en place d'un tel système car la TyrRS est dépourvue d'un système d'édition et présente une spécificité d'espèce qui empêche naturellement les charges croisées (**cf. D.**). La pénétration de l'acide aminé non-naturel dans la cellule est une contrainte supplémentaire de cette technique. Néanmoins, il est actuellement possible d'incorporer des acides aminés non-naturels dans *E. coli*, *S. cerevisiae*, des cellules d'insectes et de mammifères (revues par Xie et Schultz, 2006 ; Liu *et al.*, 2007).

Une technique différente pour l'incorporation d'acides aminés non-naturels a été mise au point par Suga (Lee et al., 2000). Elle consiste à utiliser un ribozyme nommé "flexizyme" sélectionné par un système d'évolution *in vitro* qui aminoacyle spécifiquement des ARNt avec

divers acides aminés non naturels (Murakami et al., 2006). La spécificité de reconnaissance de l'ARNt est assurée par la complémentarité entre le ribozyme et les 8 derniers nucléotides (résidus 66 à 72) de la branche acceptrice de l'ARNt (Ramaswamy et al., 2004). Ces ARNt peuvent ensuite être utilisés pour l'incorporation de l'acide aminé non naturel dans un système de traduction *in vitro*.

## 5. Systèmes d'aminacylation déviants

La règle de l'unicité des systèmes d'aminacylation stipule qu'à chaque acide aminé doit correspondre une aaRS et une famille d'ARNt isoaccepteurs, chaque organisme étant ainsi pourvu d'un jeu de 20 synthétases (Becker et Kern, 1998). Pourtant l'analyse des génomes révèle que seule une faible proportion d'organismes possède un jeu complet de synthétases. Certaines aaRS sont dupliquées, d'autres absentes, nécessitant des mécanismes détournés de synthèse des aminoacyl-ARNt correspondants (Cathopoulis et al., 2007).

### 5.1. Les voies de formation indirectes des aminoacyl-ARNt

De nombreuses bactéries et archaea dépourvues d'AsnRS et/ou de GlnRS ont recours à une voie de formation indirecte des aminoacyl-ARNt manquants. La formation de l'Asn-ARNt<sup>Asn</sup> est assurée grâce à une AspRS non discriminante qui charge l'acide aspartique sur l'ARNt<sup>Asn</sup> avant d'être convertie en asparagine par une amidotransférase (Becker et Kern, 1998). De même, la formation du Gln-ARNt<sup>Gln</sup> fait intervenir une GluRS à spécificité relâchée puis une amidotransférase (Tumbula et al., 2000). Les archaea méthanogènes sont dépourvues de CysRS, mais synthétisent le Cys-ARNt<sup>Cys</sup> par un mécanisme alternatif en deux étapes. Tout d'abord, un paralogue tétramérique de la sous-unité  $\beta$  de la PheRS récemment nommée SepRS transfère la *O*-phosphosérine (Sep) sur l'ARNt<sup>Cys</sup>. Le Sep-ARNt<sup>Cys</sup> est ensuite converti en Cys-ARNt<sup>Cys</sup> par la Sep-ARNt:Cys-ARNt synthase (SepCysS) (Fukunaga et Yokoyama, 2007b).

## 5.2. Les 21<sup>ème</sup> et 22<sup>ème</sup> acides aminés

Pendant plus d'un demi-siècle il était clairement établi que le code génétique spécifiait 20 acides aminés. Cette règle a pourtant été réfutée par la découverte d'un 21<sup>ème</sup> puis d'un 22<sup>ème</sup> acide aminé, la sélénocystéine et la pyrrolysine, respectivement. Ces deux acides aminés jouent un rôle essentiel dans les protéines qui les renferment en participant activement à la formation du site actif et à la catalyse.

La sélénocystéine, 21<sup>ème</sup> acide aminé naturel, est présente dans des protéines des trois règnes du vivant, appelées sélénoprotéines. Elles remplissent différentes fonctions, de l'oxydoréduction de métabolites dans les bactéries à la maturation de l'hormone thyroïde. Pourtant aucune SecRS n'a été identifiée. La formation du Sec-ARNt<sup>Sec</sup> s'effectue par une voie ARNt-dépendante en deux étapes. La sérine est d'abord chargée sur l'ARNt<sup>Sec</sup> par la SerRS puis convertie en Sec-ARNt<sup>Sec</sup> par la sélénocystéine synthase. Cet aminoacyl-ARNt est pris en charge par un facteur d'élongation spécifique SelB qui l'achemine au ribosome où il va décoder un codon stop (Allmang et Krol, 2006).

La pyrrolysine, 22<sup>ème</sup> acide aminé, dérive de la lysine et est présente dans certaines protéines du métabolisme du méthane d'archaea méthanogènes du genre *Methanosarcina* mais aussi dans une eubactérie gram positif *Desulfitobacterium hafniense*. La synthèse du Pyl-ARNt<sup>Pyl</sup> fait intervenir une synthétase de classe II, la PylRS et un ARNt particulier, l'ARNt<sup>Pyl</sup>, présentant un anticodon CUA complémentaire au codon stop UAG (ambre). La PylRS active la pyrrolysine avec l'ATP pour former le Pyl-AMP avant de transférer la pyrrolysine sur l'ARNt<sup>Pyl</sup>. Le Pyl-ARNt<sup>Pyl</sup>, contrairement au système Sec, est acheminé au ribosome par le facteur d'élongation classique EF-Tu (Théobald-Dietrich et al., 2004).

## D. Les systèmes de tyrosylation

Les systèmes d'aminacylation spécifiques de la tyrosine font partie des systèmes les plus étudiés au niveau biochimique, enzymologique et structural. De nombreuses données fonctionnelles et structurales ont été accumulées depuis plus d'un demi-siècle avec des études sur des organismes de trois règnes du vivant et même récemment chez un virus. L'état actuel des connaissances sur ce système sera présenté au travers d'un article de revue comportant trois grands volets : il fait le point sur (i) les jeux d'identité des ARNt<sup>Tyr</sup>, (ii) les particularités structurales des TyrRS des différents règnes du vivant et (iii) l'histoire évolutive de ce système d'aminacylation.

Article de revue n°1

### **Évolution des systèmes d'aminacylation ARNt<sup>Tyr</sup>/TyrRS**

Luc Bonnefond, Richard Giegé, Joëlle Rudinger-Thirion

**2005**

Biochimie, **87** : 873–883

Les jeux d'identité des ARNt qui assurent la fidélité de la traduction sont globalement conservés au cours de l'évolution. Les TyrRS font exception à cette règle car elles présentent une spécificité d'espèces pour la reconnaissance de l'ARNt<sup>Tyr</sup>. Cette discrimination repose sur la présence d'une paire de bases localisée à l'extrémité de la branche acceptrice de l'ARNt<sup>Tyr</sup>. Il s'agit d'une paire G1-C72 chez les bactéries qui est invariablement remplacée par une paire C1-G72 chez les archaea et les eucaryotes. En plus du rôle clé de cette paire de bases 1-72, la base discriminatrice A73, le triplet anticodon et la grande région variable (présente chez les ARNt<sup>Tyr</sup> eubactériens mais absente chez ceux d'archaea et d'eucaryotes) contribuent également à la réaction d'aminacylation, à des degrés variables. Les structures cristallographiques de deux complexes ARNt<sup>Tyr</sup>/TyrRS ont révélé des modes d'interactions avec l'ARNt<sup>Tyr</sup> différents et en accord avec la spécificité d'espèces. Les récentes études fonctionnelles sur le système ARNt<sup>Tyr</sup>/TyrRS mitochondrial humain ont mis au jour des différences significatives par rapport aux règles de tyrosylation établies. Ces différences sont discutées à la lumière des connaissances actuelles sur les TyrRS.



# Evolution of the tRNA<sup>Tyr</sup>/TyrRS aminoacylation systems

Luc Bonnefond, Richard Giegé, Joëlle Rudinger-Thirion \*

*Département 'Mécanismes et Macromolécules de la Synthèse Protéique et Cristallogénèse', UPR 9002, Institut de Biologie Moléculaire et Cellulaire du CNRS, 15, rue René Descartes, 67084 Strasbourg cedex, France*

Received 28 December 2004; accepted 17 March 2005

Available online 08 April 2005

## Abstract

The tRNA identity rules ensuring fidelity of translation are globally conserved throughout evolution except for tyrosyl-tRNA synthetases (TyrRSs) that display species-specific tRNA recognition. This discrimination originates from the presence of a conserved identity pair, G1–C72, located at the top of the acceptor stem of tRNA<sup>Tyr</sup> from eubacteria that is invariably replaced by an unusual C1–G72 pair in archæal and eubacterial tRNA<sup>Tyr</sup>. In addition to the key role of pair 1–72 in tyrosylation, discriminator base A73, the anticodon triplet and the large variable region (present in eubacterial tRNA<sup>Tyr</sup> but not found in eukaryal tRNA<sup>Tyr</sup>) contribute to tyrosylation with variable strengths. Crystallographic structures of two tRNA<sup>Tyr</sup>/TyrRS complexes revealed different interaction modes in accordance with the phylum-specificity. Recent functional studies on the human mitochondrial tRNA<sup>Tyr</sup>/TyrRS system indicates strong deviations from the canonical tyrosylation rules. These differences are discussed in the light of the present knowledge on TyrRSs.

© 2005 Elsevier SAS. All rights reserved.

*Keywords:* Mitochondria; Phylogeny; Species-specificity; Tyrosine identity

## Contents

1. Introduction .....	874
2. Peculiarities in the tyrosine identity rules .....	874
2.1. Structurally distinct tRNA <sup>Tyr</sup> species throughout evolution .....	874
2.2. Anticodon and A73 mainly determine eubacterial tyrosine identity .....	874
2.3. The first base-pair C1–G72 is the major tyrosine identity in eukarya and archæa .....	874
2.4. Base-pair N1–N72 prevents cross-aminoacylation .....	875
2.5. Structural and functional peculiarities of human mt-tRNA <sup>Tyr</sup> .....	876
2.6. tRNA mimics are substrates for TyrRSs .....	876
3. Tyrosyl-tRNA synthetases are idiosyncrasic synthetases .....	877
3.1. TyrRSs have various architectural organizations in the different domains of life .....	877
3.2. TyrRS structures in their free and complexed states .....	877
3.3. Diverse tRNA <sup>Tyr</sup> recognition modes .....	878
3.4. How overcoming the species-specificity? .....	879
3.5. Towards a novel mode of recognition in mitochondria? .....	879

*Abbreviations:* aaRS, aminoacyl-tRNA synthetase with aa for amino acid (for individual aaRSs, aa is given in the three letter code, e.g., TyrRS for tyrosyl-tRNA synthetase); BMV, brome mosaic virus; CP1, connective peptide 1; EMAP II, endothelial-monocyte-activating polypeptide II; IL-8, interleukine 8; mt, mitochondrial (e.g. mt-TyrRS or mt-tRNA<sup>Tyr</sup> for mitochondrial TyrRS or tRNA<sup>Tyr</sup>).

\* Corresponding author. Tel.: +33 3 88 41 70 57; fax:+33 3 88 60 22 18.

*E-mail address:* [J.Rudinger@ibmc.u-strasbg.fr](mailto:J.Rudinger@ibmc.u-strasbg.fr) (J. Rudinger-Thirion).

4. Considerations on evolution .....	879
4.1. The origin of tyrosine identity .....	879
4.2. Evolution of TyrRSs .....	879
4.3. Alternative functions for TyrRSs .....	880
4.4. The tyrosine system as a tool to expand the genetic code .....	881
Acknowledgements .....	881
References .....	881

## 1. Introduction

Aminoacyl-tRNA synthetases (aaRSs) are ancient enzymes, probably among the first to emerge from the RNA world [1–3]. Each of the 20 aaRSs attaches specifically an amino acid to the 3'-end of its corresponding tRNA in aminoacylation reactions whereby the link between the genetic code of four letters and the 20 amino acids of the proteins is made. The aaRSs are ranked in two classes [4,5] (themselves divided in subclasses [6,7]) that evolved from unrelated common ancestors corresponding to the catalytic core of the today synthetases [8]. Class I aaRSs share a common active site domain with a Rossmann-fold composed of alternative  $\beta$ -strands and  $\alpha$ -helices whereas the active site of class II aaRSs are framed by a seven-stranded  $\beta$ -sheet with flanking  $\alpha$ -helices. This partition, generally conserved throughout evolution, is further correlated with the type of interaction made between the synthetase and the corresponding tRNA. Whereas class I aaRSs approach their tRNAs via the minor groove side of the acceptor arm, class II aaRSs interact with the opposite minor groove side. As a consequence, attachment of the amino acid on the tRNA occurs at the 2'-OH or the 3'-OH group of the terminal adenosine.

Contemporary rules governing the specific recognition of tRNAs by their homologous synthetases have been deciphered in eubacteria and most eukarya, and constitute a second genetic code (reviewed in [9,10]). A general trend is the global conservation throughout evolution of the identity set for a given amino acid. However, the tyrosine aminoacylation system escapes this rule. Both tRNA<sup>Tyr</sup> species and TyrRSs have developed specific features that prevent cross-tyrosylation reactions in systems encompassing tRNA<sup>Tyr</sup> and TyrRS from different domains of life (see [11] for a review on the enzymology of TyrRSs). In our review, the peculiarities of the tyrosine systems, at both structural and functional levels are emphasized. The main differences between eukaryal/archæal and eubacterial tyrosylation systems will be recapitulated in the light of the available crystal structures. It will be shown how benefit can be taken from these differences to expand the genetic code. Finally, accent will be given to the unforeseen functional behavior of human mt-TyrRS as well as to the cellular processes, different from protein synthesis, in which participates TyrRS or tRNA<sup>Tyr</sup> mimics.

## 2. Peculiarities in the tyrosine identity rules

### 2.1. Structurally distinct tRNA<sup>Tyr</sup> species throughout evolution

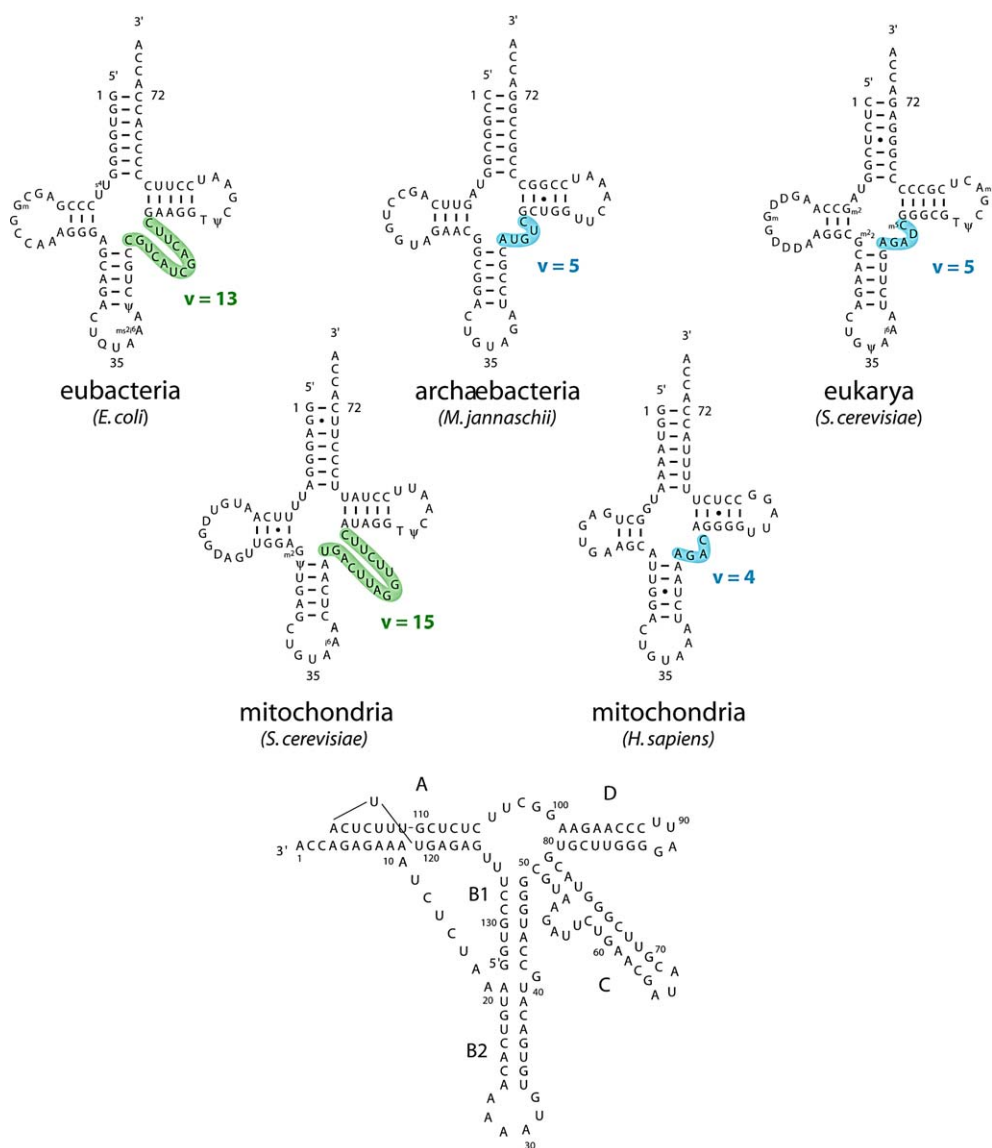
Most tRNAs fold into the standard cloverleaf secondary structure and are classified according to the length of their variable region [12]. While the majority of tRNAs belong to class 1 with four to five nucleotides in their variable loop, four of them (tRNA<sup>Leu</sup>, tRNA<sup>Tyr</sup>, tRNA<sup>Ser</sup> and tRNA<sup>Sec</sup>) are members of class 2 with more than 10 nucleotides in this region. Tyrosine specific tRNAs are unique in that their variable region is large in eubacteria whereas it is short in eukarya and archæa. Noteworthy are also mitochondrial (mt) tRNA<sup>Tyr</sup> molecules that share a four-nucleotide long variable region when from higher eukarya but belong to class 2 in lower eukarya. Thus, tRNA<sup>Tyr</sup> species constitute a unique example where a secondary folding feature is not conserved throughout evolution (Fig. 1).

### 2.2. Anticodon and A73 mainly determine eubacterial tyrosine identity

Study of amber suppressor tRNA<sup>Tyr</sup> from *Escherichia coli* revealed in the 70s the importance for tyrosylation of the terminal base-pair next to the CCA accepting triplet [13,14]. Later, tyrosine identity in eubacteria was investigated using in vitro and in vivo methods. The role of the phylogenetically conserved discriminator base A73 and of anticodon residues G34 and U35 in identity was demonstrated by the partial to complete loss of tyrosylation ability of the tRNA after mutation of these positions [15–18]. A characteristic of eubacterial tRNA<sup>Tyr</sup> is the large variable region whose orientation is dependent on the number of unpaired nucleotides at its 3'-end. This orientation rather than the sequence of the variable region is important for tyrosine identity [18]. A less drastic involvement in identity was found for the G1–C72 pair via the use of *E. coli* tRNA<sup>Tyr</sup> microhelices that show their tyrosylation efficiency reduced after mutating this pair [19]. All these data are summarized in Table 1.

### 2.3. The first base-pair C1–G72 is the major tyrosine identity in eukarya and archæa

First attempts to decipher tyrosine identity in yeast were performed by chemical microsurgery of the tRNA [20,21].



### BMV tRNA-like structure

Fig. 1. Secondary structures of a selection of tRNA<sup>Tyr</sup> species and of the tyrosylatable domain of BMV RNA. For the tRNAs, residues are numbered according to [12]. Sequence is either that of native tRNA<sup>Tyr</sup> (*E. coli* and cytosolic *S. cerevisiae*) or of transcripts (*M. jannaschii* and mitochondrial species). The variable region (v) is emphasized in green when large ( $v > 10$  nt) and in blue when short (4–5 nt). Note the numbering of the BMV tRNA-like structure that starts from the 3'-end.

Using this technology, it could be demonstrated that anticodon loop nucleotides are required for optimal aminoacylation, in particular residue G34. A more central role for base-pair G1–C72 and discriminator residue A73 came from in vivo experiments performed with mutant suppressor tRNAs in yeast [22]. In addition, systematic determination of kinetic parameters of yeast tRNA<sup>Tyr</sup> variants mutated at the strategic positions confirmed the importance of these latter elements in tyrosylation [23]. Further in vitro data generalized the crucial role of pair C1–G72 since charging of yeast microhelices by TyrRS from the eukaryal pathogen *Pneumocystis carinii* is abolished after mutating C1–G72 into G1–C72 [19].

The aforementioned identity determinants are also effective in the archæon *Methanococcus jannaschii* with however a weaker participation of the anticodon residues as compared

to the situation in yeast [24]. Due to the reduced contribution of anticodon, minihelices derived from *M. jannaschii* tRNA<sup>Tyr</sup> (Fig. 1) could be charged [25]. Their activity follows the tyrosylation rules since mutation of base-pair C1–G72 in the reverse pair impairs tyrosylation [24]. Data are recapitulated in Table 1.

#### 2.4. Base-pair N1–N72 prevents cross-aminoacylation

All eubacterial tyrosine tRNAs have a G1–C72 base-pair while all eukaryal and archæal species share the reverse C1–G72 base-pair, a pair uncommon in other tRNAs [12]. Since N1–N72 pairs are different and are required for tyrosylation in both domains of life, they appear to be the main species-specific determinants for tyrosylation. Indeed, a strong



Table 1  
Pattern and strength of the tyrosine identity elements for specifying tRNA tyrosylation in the different domains of life and in human mitochondria

Organism	tRNA <sup>Tyr</sup> nucleotides	Involvement	References
Eubacteria			
<i>E. coli</i>	A73	++	[16–18]
	G1–C72	+	[16,18,19]
	G34	+	[15]
	U35	++	[18]
	A36		No data available
	Variable region <sup>a</sup>	++	[18]
Archæobacteria			
<i>M. jannaschii</i>	A73	++++	[24]
	C1–G72	++++	[24]
	G34	++	[24]
	U35	+	[24,25]
	A36	+	[24]
Eukarya			
<i>S. cerevisiae</i>	A73	++++	[22,23]
	C1–G72	+++	[22,23]
	G34	+++	[20,21,23]
	U35	++	[20,21,23]
	A36	++	[23]
<i>P. carinii</i>	C1–G72	+++	[19]
Mitochondria			
<i>H. sapiens</i>	A73	+++	[29]
	G34	+++	Present work
	U35	+++	Present work
	A36	+++	Present work

Involvement of each element (estimated by the  $k_{cat}/K_m$  effect) ranges from (++++) for strong to (+) for weak. Tyrosylation data of human mt-tRNA<sup>Tyr</sup> were obtained under the experimental conditions described by [44].

<sup>a</sup> Variable region is recognized for its structure rather than its sequence.

and unusual phylogenetic barrier exists preventing cross-tyrosylation of eubacterial tRNA<sup>Tyr</sup> by eukaryal TyrRSs and vice-versa [26,27]. Note that this behavior is unique among aminoacylation systems since identity rules are generally conserved throughout evolution [9].

### 2.5. Structural and functional peculiarities of human mt-tRNA<sup>Tyr</sup>

Most mitochondrial tRNA<sup>Tyr</sup> species possess a G1–C72 base-pair as do eubacterial tRNA<sup>Tyr</sup> species [12]. Concerning the variable region, mt-tRNA<sup>Tyr</sup> is related to eukaryal tRNA<sup>Tyr</sup> with only five nucleotides in this part of the molecule (Fig. 1). The situation differs in mitochondria from lower eukarya where the tyrosine tRNAs contain a long variable region. We recently kept our attention on human mt-tRNA<sup>Tyr</sup>, representative of mt-tRNAs from higher eukarya. Although this molecule has a classical cloverleaf structure (Fig. 1), its 3D folding into the standard L-shaped structure does not occur via the classical long-range interactions since the required nucleotides are lacking in mt-tRNA<sup>Tyr</sup>, as usual for mammalian mt-tRNAs [28]. Remarkably, this human mitochondrial tyrosylation system behaves differently from those studied so far [29]. Indeed, mt-TyrRS aminoacylates not only its homologous tRNA<sup>Tyr</sup> transcript with G1–C72 but also the

C1–G72 variant, as well as heterologous tRNA<sup>Tyr</sup> molecules disregarding the nature of the N1–N72 pair. This indicates that pair G1–C72 in human mt-tRNA<sup>Tyr</sup> is not part of the identity set, an unexpected conclusion we will discuss in more details later. Further, mutating the discriminator and the anticodon bases revealed a strong involvement of these elements in tyrosine identity with losses of catalytic efficiency of 5000 for the G73 mutation [29] and ranking from 2000 to 6000 for those at anticodon positions (Table 1). The fact that C1–G72 does not belong to the tyrosine identity set seems to be compensated by an enhanced involvement in identity of discriminator and anticodon residues.

### 2.6. tRNA mimics are substrates for TyrRSs

Natural tRNA mimics can be substrates of synthetases and the signals specifying their recognition by these enzymes are mostly at similar locations as in canonical tRNA [30,31]. Two mimics of tRNA<sup>Tyr</sup> will illustrate this view. The first one that can be efficiently tyrosylated by yeast and other eukaryal TyrRSs corresponds to the 3'-end of brome mosaic virus (BMV) RNA [32–34]. Residues specifying its tyrosylation by the yeast enzyme are contained within the 134 nucleotides of the 3'-end of the so-called tRNA-like structure of this viral RNA (Fig. 1). Notice the pseudo-knotted architecture of the amino acid acceptor branch of the BMV RNA. A mutational analysis revealed the determinants conferring the tyrosine identity to this structure. They are residues A4 and base-pair C116–G5 homologous to A73 and base-pair C1–G72 in eukaryal tRNA<sup>Tyr</sup>. An additional stem-loop domain present in the intricate fold of BMV RNA acts as a non-specific anchor for enhancing the affinity of the tRNA-like structure for TyrRS [35]. However, no functional equivalent of the tyrosine anticodon was found in the tRNA-like structure, reducing the tyrosine identity set to the sole residues located in the acceptor branch. Thus, the BMV tRNA-like structure follows the same rules for tyrosylation than canonical tRNA<sup>Tyr</sup> as far as the accepting branch is concerned. In other words this molecule behaves as an aminoacylatable minihelix.

The second example concerns group I introns of *Neurospora crassa* whose catalytic core interact with the mitochondrial TyrRS (CYT-18 protein) [36–38]. This CYT-18 protein contains two small insertions not found in non-splicing eubacterial enzymes [38]. Probing experiments combined with modeling revealed that group I introns contain a tRNA-like structural motif similar to the homologous mt-tRNA<sup>Tyr</sup> including the presence of an equivalent of the long variable arm [39]. Here, no aminoacylation takes place but the interaction between both molecules constitutes a key step in the group I intron-folding pathway. Further, probing and modeling showed that the insertions in CYT-18 are essential for splicing [38].

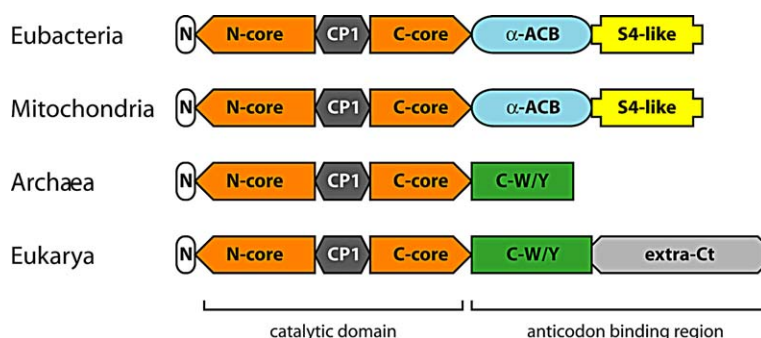


Fig. 2. Structural organization of TyrRSs in the three domains of life and in mitochondria. Each structural domain of TyrRSs is given in a specific color. The abbreviations are: (N) for N-terminal motif, (N-core) and (C-core) for N- and C-parts of the catalytic domain, (CP1) for connective peptide, ( $\alpha$ -ACB) for  $\alpha$ -helical anticodon-binding domain, (C-W/Y) for C-terminal domain homologue to TrpRS, (S4-like) for S4-like domain and (extra-Ct) for additional C-terminal domain (adapted from [42]).

### 3. Tyrosyl-tRNA synthetases are idiosyncrasic synthetases

#### 3.1. TyrRSs have various architectural organizations in the different domains of life

All TyrRSs are homodimers with their N-terminal domain containing an active site based on a parallel  $\beta$ -sheet nucleotide binding fold or Rossmann-fold [4,5], typical of class I synthetases. However, the two class I signature sequences HIGH and KMSKS are not strictly conserved and can be replaced by sequences as different as HAYL and KFGKT. The catalytic domain of TyrRSs contains ~230 amino acids and is conserved in structure (Fig. 2). It is parted by a short peptide insertion (called CP1) of ~50 amino acids. In contrast to the N-terminal half, the C-terminal moiety varies considerably in sequence, length and organization. It is composed in all eubacterial TyrRSs by a  $\alpha$ -helical domain ( $\alpha$ -ACB) of ~100 amino acids, followed by a domain (called S4-like) that shares high homology with the C-terminal domain of ribosomal protein S4 [40]. Variability in their C-terminal parts leads to the ranking of eubacterial TyrRSs into two sub-groups [41] (see below for more details).

Archæal TyrRSs are peculiar in that their C-terminal moieties are significantly shorter than those of TyrRSs from the other domains of life [25]. They share strong homologies with the C-terminal domain of TrpRSs. This domain called C-W/Y [42], different from the  $\alpha$ -helical domain found in eubacteria, is present in all eukarya and is followed in higher eukarya by an extension of ~160 amino acids. In the particular case of the human TyrRS, this extension shares 51% sequence identity with the human cytokine EMAP II [26].

As for other mitochondrial proteins, mt-TyrRSs likely originate from a TyrRS of eubacterial origin, as a result of gene transfer from a proto-mitochondrion [43]. It is thus understandable that mt-TyrRSs share only little sequence identity with their cytosolic relatives but resembles more to the eubacterial TyrRSs, in particular for their modular organization with both  $\alpha$ -ACB and S4-like domains (Fig. 2). This is accurately the case for human mt-TyrRS that does not align with its cytosolic counterpart but shares up to 40% identity with *Bacillus stearothermophilus* TyrRS [44].

#### 3.2. TyrRS structures in their free and complexed states

*B. stearothermophilus* TyrRS was the first synthetase whose structure has been solved at high resolution (2.3 Å) alone [45] and in complex with tyrosine [46], tyrosyl-adenylate and tyrosinyl-adenylate [47]. More recently, the structures of *Staphylococcus aureus* TyrRS [48] and of truncated versions, deprived of the C-terminal domain, of *E. coli* [49] and mitochondrial *N. crassa* [38] TyrRSs were solved. The eubacterial TyrRSs and the related mt-TyrRS are similarly organized as schematized in Fig. 2. However, with native enzymes the C-terminal S4-like domain is not seen in the crystal structures as a result of structural disorder. The domain, as such, is not disordered since its structure was solved recently by nuclear magnetic resonance [40] and revealed a fold similar to that found in ribosomal protein S4 [50]. The CP1 domain within the Rossmann-fold, generally involved in editing in other class I synthetases [51], is small and forms here the dimer interface. A breakthrough came with the resolution of the TyrRS/tRNA<sup>Tyr</sup> complex of *Thermus thermophilus* [52]. It offers a view of the complete synthetase organization, including the C-terminal domain, and permits to visualize the interaction with the cognate tRNA<sup>Tyr</sup> [52]. This crystal structure confirms the cross-subunit binding of tRNA<sup>Tyr</sup> as first proposed by Bedouelle on the basis of an extensive mutational analysis [53,54]. As seen in Fig. 3A, the acceptor stem of tRNA<sup>Tyr</sup> interacts with the catalytic domain of one monomer, whereas the anticodon stem is sandwiched between the  $\alpha$ -helical and C-terminal domains of the other monomer. Notice that the C-terminal domain binds to the characteristic long variable stem-loop region of the tRNA. Further, as already suggested by numerous biochemical data [24,55–57], TyrRS (a class I synthetase) approaches its cognate tRNA<sup>Tyr</sup> from the acceptor stem and variable loop major groove side, a feature typical of class II synthetases. This unusual interaction scheme is consistent with the dimeric nature of the TyrRS.

In eukarya and the closely related archæa, structures of a truncated free TyrRS and of a tRNA<sup>Tyr</sup>/TyrRS complex were solved. Crystal structure of the human minimalist TyrRS deprived of the EMAP II-like domain [58] compares well to

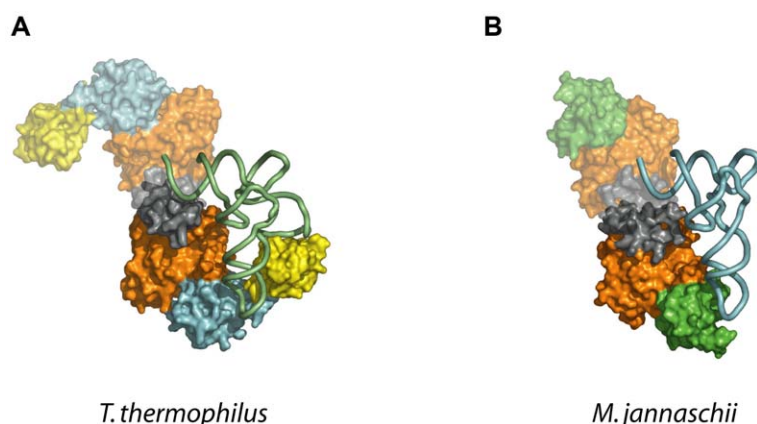


Fig. 3. Crystal structures of eubacterial and eukaryal-like TyrRS/tRNA<sup>Tyr</sup> complexes. Structures of (A) *T. thermophilus* [52] and (B) *M. jannaschii* [60] TyrRS complexed to homologous tRNA<sup>Tyr</sup>. The different domains of TyrRSs are colored according to the color code of Fig. 2. For better visualization, only one of the two tRNA<sup>Tyr</sup> molecules is depicted and one subunit of the homodimeric TyrRS is in attenuated colors. Drawings were done using PyMOL [91].

that of the archæal TyrRS from *M. jannaschii*. Interestingly, the folding of the anticodon recognition domain diverges from that of eubacterial counterparts but shares homology with eubacterial TrpRS [59]. Noteworthy the crystal structure of the TyrRS/tRNA<sup>Tyr</sup>/L-tyrosine complex from *M. jannaschii* extends to eukarya the unexpected class II mode of tRNA recognition of a class I synthetase [60] (Fig. 3B). Finally, the small C-terminal domain of the complexed TyrRS is structurally different from that of eubacterial TyrRSs although contacting the same anticodon residues in tRNA<sup>Tyr</sup>. A further difference concerns the absence of interaction between the short variable region of tRNA<sup>Tyr</sup> with *M. jannaschii* TyrRS [60], as clearly seen in Fig. 3B.

### 3.3. Diverse tRNA<sup>Tyr</sup> recognition modes

The N1–N72 base-pair determines the species-specificity for tyrosylation at the tRNA<sup>Tyr</sup> level [19]. Its presence in the G–C or C–G version is mainly responsible for the absence of cross-tyrosylation in heterologous eubacterial/eukaryal tRNA<sup>Tyr</sup>/TyrRS systems. Resolution of both *T. thermophilus* and *M. jannaschii* complexes [52,60] brought to light different acceptor stem recognition patterns in line with the functional data. In the eubacterial complex, two regions of the catalytic domain contact the acceptor stem of tRNA<sup>Tyr</sup>. They correspond to amino acids 148–161 and 198–211, previously identified as clusters 1 and 2 in *B. stearrowthermophilus* TyrRS [16,61]. Residues G1, C72 and A73 contact directly amino acids of the homologous synthetase [52]. In the archæal complex, these two clusters correspond to amino acids 126–139 and 174–187. Here, the geometry of the acceptor stem is disturbed with the C1 base twisted and forming H-bonds with several amino acids, whereas A73 is unstacked and out of the helical continuity of the acceptor stem. Specific interactions involve also residues G72 and A73 of tRNA<sup>Tyr</sup> [60]. In summary, in both complexes N1–N72 and A73 interact with the corresponding TyrRS but via vastly different amino acid side chains [62]. These differences in recognition pattern constitute the basis for the species-specificity (Fig. 4).

Anticodon residues G34, U35, A36 of tRNA<sup>Tyr</sup> are part of the tyrosine identity set but act at different strengths depending on the domain of life (Table 1). They contact the C-terminal part of TyrRSs in the complexes. However, due to the shape differences of the synthetases, the anticodon bases are recognized in a completely different manner. In the *T. thermophilus* complex, the triplet adopts a novel conformation and base specific interactions occur with G34 and U35 and not with A36, involved in van der Waal's contacts. The smaller TyrRS from *M. jannaschii* interacts specifically with the flipped out G34 residue, the other two anticodon bases having less specific interactions with the synthetase. Details about the amino acids interacting with the major tyrosine identity residues and their phylogenetic conservation are given in Fig. 4. Altogether the structural data are in line

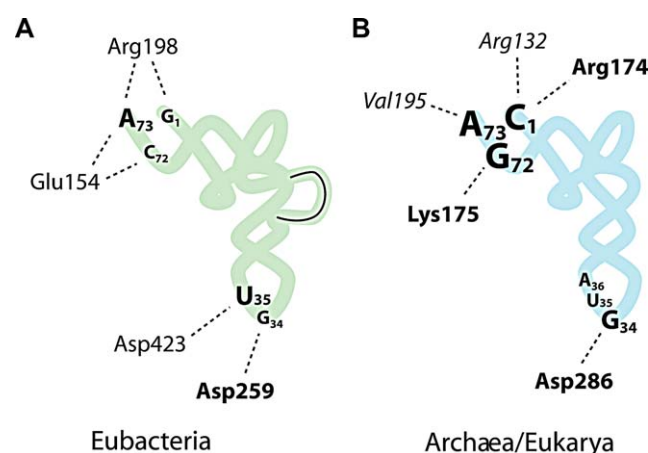


Fig. 4. Schematic view of the specific interactions between tRNA<sup>Tyr</sup> identity elements and TyrRSs. Based on crystallographic data [52,60], the amino acids of TyrRSs that contact directly the tyrosine identity nucleotides of (A) *T. thermophilus* and (B) *M. jannaschii* tRNA<sup>Tyr</sup> are indicated. Amino acids in bold are those conserved in TyrRSs from eubacteria or archæa/eukarya; those in italics are not strictly conserved but are of same nature and those in standard letters are not conserved [29]. The font size of the tRNA<sup>Tyr</sup> identity nucleotides is proportional to their effect on tyrosylation efficiency (expressed in losses of  $k_{cat}/K_M$  values) as determined (A) in *E. coli* [15–19] and (B) in *M. jannaschii* [24,25]. The long variable arm typical of eubacterial tRNA<sup>Tyr</sup> is lined since it is part of the identity set in eubacteria [18].



with the mutational analysis indicating that tRNA mutated at contact positions have decreased tyrosylation activity (Table 1).

### 3.4. How overcoming the species-specificity?

Engineering either tRNA<sup>Tyr</sup> or TyrRS permitted to surmount the strict species-specificity. Concerning tRNA<sup>Tyr</sup>, specificity swap was achieved by replacing G1–C72 by C1–G72 in *E. coli*. As anticipated, the transplanted molecule became an efficient substrate for either yeast [23] or *M. jannaschii* [24] TyrRSs. Likewise, transplantation of a small peptide corresponding to the CP1 domain, from the human into the *E. coli* TyrRS, enabled this latter to tyrosylate human tRNA<sup>Tyr</sup> and conversely [27]. This inversion of species-specificity emphasizes the preponderant role played by the amino acids cluster within the CP1 domain for N1–N72 recognition. From what discussed above, it can be inferred that the sole analysis of a new TyrRS sequence should be sufficient to predict its functional behavior towards tRNA<sup>Tyr</sup> species. Following these lines, the foreseen species-specificity of *Mycobacterium tuberculosis* TyrRS was experimentally verified [61].

One can question about the origin of these differences in the tyrosine system. A first explanation states that the sequence patterns in eubacterial and archæal/eukaryal clusters are the result of co-variations of the enzyme structure that has adapted to changes in tRNA [61].

### 3.5. Towards a novel mode of recognition in mitochondria?

Only few mitochondrial TyrRS sequences have been annotated and their corresponding enzymes studied, namely those of *N. crassa* [36], *Saccharomyces cerevisiae* [63] and *Podospora anserina* [64] for lower eukarya, and recently that of *Homo sapiens* for higher eukarya [44]. Their corresponding mt-tRNAs have a G1–C72 pair as in eubacterial tRNA<sup>Tyr</sup> species, except *N. crassa* mt-tRNA<sup>Tyr</sup> that shares an unusual G1–U72 pair [12]. Analyses of human mt-TyrRS, particularly in clusters 1 and 2, emphasize new sequence characteristics. Indeed, this synthetase is a mosaic with, as expected, eubacterial and mitochondrial sequence features but also archæal marks [29]. Thus, it was not possible to predict the recognition mode with the 1–72 pair, and indeed an unexpected functionality was found, namely the ability of human mt-TyrRS to charge tRNA<sup>Tyr</sup> with either G1–C72 or C1–G72 base-pair [29]. This property might be generalized to other mt-TyrRSs from higher eukarya (e.g. *Tetraodon nigroviridis*, *Ratus norvegicus*) since they share strong sequence similarities with human mt-TyrRS (data not shown). Finally, although human mt-TyrRS sequence reveals ~40% similarity with mt-TyrRSs from lower eukarya, these latter do not contain the aforementioned archæal marks. Nevertheless, sequence similarities with eubacterial TyrRSs suggest an interaction scheme close to that found in the *T. thermophilus* system.

## 4. Considerations on evolution

### 4.1. The origin of tyrosine identity

The catalytic domain of aaRSs, considered as the primordial synthetase [8,65,66], is generally conserved and interacts with the acceptor branch of tRNAs that contain the major identity elements. The additional and more idiosyncratic domains appeared later in evolution and interact with the tRNA anticodon branch. This view is well illustrated in the case of *M. jannaschii* TyrRS that presents a small anticodon-binding domain only poorly interacting with the anticodon of tRNA<sup>Tyr</sup>. Here, the strongest interactions take place between the catalytic domain of the synthetase and the acceptor stem of the tRNA [25]. This explains the ability of the archæal enzyme to tyrosylate minihelices that recapitulate the amino acid acceptor branch of *M. jannaschii* tRNA<sup>Tyr</sup> [24]. Thus, the archæal TyrRS is representative of an early synthetase in contrast to eubacterial or mitochondrial TyrRSs that share larger C-terminal domains interacting more strongly with the anticodon residues.

Inspection of modern tyrosine systems suggests a correlation between the nature of the identity residues in the tRNA<sup>Tyr</sup> acceptor stem with the size of the variable region. Striking is the concomitant presence of a large variable region together with a G1–C72 pair in eubacterial/mitochondrial tRNA<sup>Tyr</sup> and the systematic replacement of these two elements by a small variable region and a C1–G72 pair in all archæal/eukaryal tRNA<sup>Tyr</sup>. This could mean that both elements evolved simultaneously. Mammalian mitochondrial tRNA<sup>Tyr</sup> species are the only exceptions since they have G1–C72 pairs and small variable regions. To rationalize these observations one can propose the following scenario. A proto-tRNA<sup>Tyr</sup> would have been a minihelix with A73 and G1–C72. Because these sequence elements are shared with many other tRNAs, the tyrosine identity of this minihelix was relaxed. Specificity was reached when additional identity signals emerged as the result of a mutual adaptation of the primitive tyrosine tRNAs and the TyrRSs. The process was not univocal but found two solutions. In eubacteria, it occurred when the tRNA acquired a large variable region and when in parallel the TyrRS learned to recognize this domain. In eukarya/archæa the situation was different since the inversion of the G1–C72 pair into the rare C1–G72 pair prevented the TyrRS to adapt to a particular RNA architecture. This would account a posteriori for a tRNA-triggered evolution of the TyrRSs that acquired their additional modules as specific answers allowing recognition of alternate architectural characteristics of their tRNA substrates, e.g. the long variable arm and/or the anticodon domain of tRNA<sup>Tyr</sup>. Such a scenario did not occur for other identities brought by class 2 tRNAs (i.e. leucine and serine specific).

### 4.2. Evolution of TyrRSs

TyrRSs belong to the same sub-group than TrpRSs in class I synthetases. This ranking finds strong support when com-

paring their sequences. Pairwise comparison of TyrRS/TrpRS sequences reveals among others that archæal/eukaryal TyrRSs and TrpRSs are closer to each other than to their eubacterial counterparts. From this observation it was postulated that both families of aaRSs diverged after the separation of the three domains of life [67]. However, an alternate analysis of sequence data supports the view that TyrRSs and TrpRSs form two separate monophyletic groups [68,69]. Although the N-terminal catalytic domains of all TyrRSs are similarly organized, the C-terminal moieties show variability between eubacteria and archæa/eukarya. The anticodon-binding domain (C-W/Y) common to archæal/eukaryal TyrRSs and TrpRSs is lost in eubacterial TyrRSs [42] and replaced by both  $\alpha$ -helical and S4-like domains (Fig. 2). This replacement is considered as an ancient event since it predates the separation of all known eubacteria. In parallel, fusion of a RNA-binding domain to the C-terminal of TyrRSs in higher eukarya occurred [70]. In fact, this extra-domain along with the IL8 characteristic ELR motif at the N-terminus is present in segmented animals whereas these two features are absent in yeast and lower eukarya (e.g. in *Caenorhabditis elegans*) [71].

Comparison of all known TyrRS sequences lead to the phylogenetic tree schematized in Fig. 5. As anticipated, mt-TyrRSs branch close to eubacterial TyrRSs. These latter are partitioned into two sub-groups, called TyrRSs and TyrRZs, with at most 25% sequence identity [41,72]. Sequence alignments between both types of enzymes show that most differences are contained within the C-terminal domain. Despite this divergence, some residues important for catalysis and substrates binding are similar and probably play the same function in both sub-groups [73]. This partition could result from a horizontal transfer of TyrRS genes between the two groups [42]. Alternatively, TyrRS genes could have duplicated early and evolved differently due to an environmental pressure [68]. However, this partition does not correlate with taxonomic criteria [42]. The tree also shows that archæal TyrRSs are parted into two sub-groups, the euryarchæotal TyrRSs giving rise to eukaryal TyrRSs, and the crenarchæotal TyrRSs from which plant-, plasmodium- and mimivirus-TyrRSs are derived. Noticeable is the newly discovered mimivirus that encodes four aaRSs including a TyrRS. In accordance with the authors that position this giant virus in an independent branching near the origin of eukarya [74], it appears that its TyrRS follows the same phylogenetic ranking with a branching originating from euryarchæota (Fig. 5).

Remarkable is the evolutionary divergence of TyrRSs where overall conserved catalytic N-terminal moieties are able to discriminate between C1–G72 and G1–C72 pairs in tRNA<sup>Tyr</sup> species. On the contrary, convergent evolution took place in the C-terminal moieties to recognize the conserved G34-U35-A36 tyrosine anticodon triplet with structurally different domains.

#### 4.3. Alternative functions for TyrRSs

The C-terminal extra-domain of several higher eukaryal aaRSs is usually involved in the assembly of multi-enzymatic

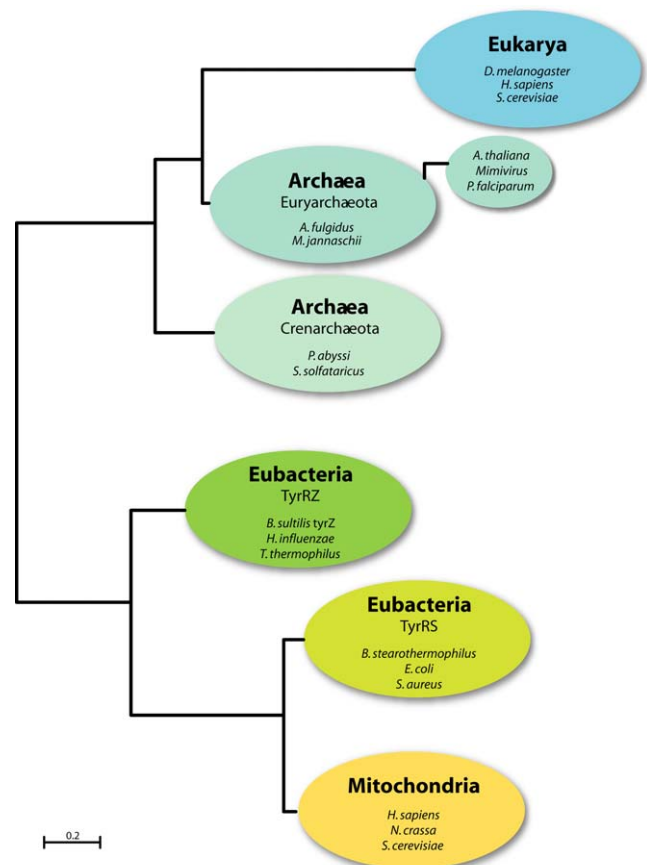


Fig. 5. Phylogenetic tree of TyrRS sequences. A set of 50 TyrRS sequences from various organisms of each phylae was used. An ungapped alignment of the catalytic domain of those TyrRSs was made using 3DCoffee [92] to build a tree with the PHYLIP package [93]. Scale bar represents amino acids substitutions per site. In each phylogenetic group, up to three representative species are indicated. Organisms listed in the tree are *Arabidopsis thaliana*, *Archaeoglobus fulgidus*, *B. stearothermophilus*, *Bacillus subtilis*, *Drosophila melanogaster*, *Haemophilus influenzae*, *H. sapiens*, *M. jannaschii*, *Mimivirus*, *N. crassa*, *S. cerevisiae*, *Plasmodium falciparum*, *Pyrococcus abyssi*, *S. aureus*, *Sulfolobus solfataricus*, *T. thermophilus*.

complexes [75]. Human cytosolic TyrRS however is not part of this complex and its C-terminal extra-domain shows no homology with the nine synthetases of the multi-enzymatic complex but, in contrast, with the human cytokine EMAP II [26]. Under apoptotic conditions, human TyrRS is split by an extracellular enzyme into two fragments [76]. As expected, the released C-terminal domain, specific of all mammalian TyrRSs, shares the expected cytokine activity. Thus, insertion of an idiosyncratic domain at the C-terminus end of the ancestral TyrRS gave a novel function. As to the N-terminal moiety of the contemporary protein that shows tyrosylation properties, it can also function as a cytokine (IL8-like) due to the presence of an ERL sequence motif within the Rossmann-fold. This motif is conserved in TyrRSs from mammals but not in the enzymes from lower eukarya [76].

Another role described for two yeast mt-TyrRSs is to participate in splicing of group I introns. Particularly well studied is the mt-TyrRS from *N. crassa* (CYT-18 protein) that charges its cognate tRNA<sup>Tyr</sup> but also recognizes (but not ami-

noacylates) a shape in group I introns that mimic *N. crassa* tRNA<sup>Tyr</sup> including its large variable region [39]. Additional small idiosyncratic extensions at the N-terminus of mt-TyrRS plays also a key role in the splicing activity [38,77]. It was proposed that, in earlier time, group I introns were able to self-splice and became later dependent of protein CYT-18 that shared the needed prerequisites. This bi-functionality was also demonstrated for mt-TyrRS from *P. anserina* [64].

Remarkably, the core of *N. crassa* mt-TyrRS participates in two distinct biological functions, aminoacylation and splicing, in contrast to the human TyrRS where the cytokine function is ensured by an appended domain unrelated with the aminoacylation function. Other surprising functions besides aminoacylation were found with TyrRSs from rabbit liver that presents a kinase activity [78] and TyrRS from *S. cerevisiae* that is involved in the maintenance of the mitochondrial genome as shown by genetic tools [79]. Further studies are needed to attribute these novel properties to specific domains of the corresponding TyrRSs.

#### 4.4. The tyrosine system as a tool to expand the genetic code

The strategy used to expand the genetic code in vivo consists in engineering synthetase/tRNA couples to produce a tRNA specifically charged with an unnatural amino acid by a synthetase, orthogonality of both macromolecular partners being the main prerequisite [80,81]. In other words, an orthogonal tRNA (often an amber suppressor with G34 mutated into C34) should be exclusively charged with a novel amino acid by an orthogonal synthetase that excludes all the tRNAs present in the cell. Incorporation of the new amino acid can then occurs at a selected target position (often amber stop codon) in a given protein. The peculiar properties of TyrRSs, namely the absence of cross-aminoacylations, the low involvement of anticodon residue 34 in identity, and the potential of TyrRSs to activate tyrosine analogs and even to transfer them on tRNA (e.g. [82] for D-tyrosine incorporation, and reviewed in [11]), have made tyrosine systems interesting candidates for engineering orthogonal tRNA/aaRS pairs for incorporation of unnatural amino acids into proteins (e.g. [60,62,80,81,83–88]). Following these lines it became possible by genetic screens to tune *M. jannaschii* TyrRS and suppressor tRNA<sup>Tyr</sup> so that new amino acids, such as O-methyl-L-tyrosine, could be incorporated into *E. coli* proteins (e.g. [83]). This strategy was enlarged for site-specific incorporation of the unnatural 3-iodo-L-tyrosine into proteins in eukaryal translation systems [89,90]. Finally, the detailed crystallographic knowledge of tRNA<sup>Tyr</sup>/TyrRS complexes that highlighted the different modes of recognition in eubacteria and eukarya, will enable on a rational basis to engineer *M. jannaschii* TyrRS so that to compensate for the lack of efficiency in tyrosylation of the homologous tRNA<sup>Tyr</sup> amber suppressor [60,62]. Use of such TyrRS variants, by enhancing the amber suppression efficiency, will than be sufficient to produce large amounts of alloproteins.

#### Acknowledgements

We thank Jean-Michel Claverie and Chantal Abergel for communicating us unpublished data on mimivirus TyrRS. This work was supported by Centre National de la Recherche Scientifique (CNRS), Université Louis Pasteur, Strasbourg, and the French Ministry for Research (ACI “BCMS”).

#### References

- [1] C. Hountondji, P. Dessen, S. Blanquet, Sequence similarities among the family of aminoacyl-tRNA synthetases, *Biochimie* 68 (1986) 1071–1078.
- [2] P. Schimmel, Aminoacyl-tRNA synthetases: general scheme of structure–function relationships in the polypeptides and recognition of transfer RNAs, *Annu. Rev. Biochem.* 56 (1987) 125–158.
- [3] D. Moras, Structural and functional relationships between aminoacyl-tRNA synthetases, *Trends Biochem. Sci.* 17 (1992) 159–164.
- [4] S. Cusack, C. Berthet-Colominas, M. Härtle, N. Nassar, R. Leberman, A second class of synthetase structure revealed by X-ray analysis of *Escherichia coli* seryl-tRNA synthetase at 2.5 Å, *Nature* 347 (1990) 249–255.
- [5] G. Eriani, M. Delarue, O. Poch, J. Gangloff, D. Moras, Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs, *Nature* 347 (1990) 203–206.
- [6] S. Cusack, Aminoacyl-tRNA synthetases, *Curr. Opin. Struct. Biol.* 7 (1997) 881–889.
- [7] L. Klipcan, M. Safro, Amino acid biogenesis, evolution of the genetic code and aminoacyl-tRNA synthetases, *J. Theor. Biol.* 228 (2004) 389–396.
- [8] P. Schimmel, R. Giegé, D. Moras, S. Yokoyama, An operational RNA code for amino acids and possible relationship to genetic code, *Proc. Natl. Acad. Sci. USA* 90 (1993) 8763–8768.
- [9] R. Giegé, M. Sissler, C. Florentz, Universal rules and idiosyncratic features in tRNA identity, *Nucleic Acids Res.* 26 (1998) 5017–5035.
- [10] P.J. Beuning, K. Musier-Forsyth, Transfer RNA recognition by aminoacyl-tRNA synthetases, *Biopolymers* 52 (1999) 1–28.
- [11] H. Bedouelle, Tyrosyl-tRNA Synthetases, in: M. Ibba, C. Francklyn, S. Cusack (Eds.), *Aminoacyl-tRNA Synthetases*, Landes Biosciences, Georgetown, TX, 2004, pp. 111–124 Chapter 12.
- [12] M. Sprinzl, K.S. Vassilenko, Compilation of tRNA sequences and sequences of tRNA genes, *Nucleic Acids Res.* 33 (2005) D139–D140.
- [13] J.D. Smith, J.E. Celis, Mutant tyrosine transfer RNA that can be charged with glutamine, *Nat. New Biol.* 243 (1973) 66–71.
- [14] J.E. Celis, M.L. Hooper, J.D. Smith, Amino acid acceptor stem of *E. coli* suppressor tRNA<sup>Tyr</sup> is a site of synthetase recognition, *Nat. New Biol.* 244 (1973) 261–264.
- [15] Y.M. Hou, P. Schimmel, Modeling with in vitro kinetic parameters for the elaboration of transfer RNA identity in vivo, *Biochem.* 28 (1989) 4942–4947.
- [16] H. Bedouelle, Recognition of tRNA<sup>Tyr</sup> by tyrosyl-tRNA synthetase, *Biochimie* 72 (1990) 589–598.
- [17] J.M. Sherman, K. Rogers, M.J. Rogers, D. Söll, Synthetase competition and tRNA context determine the in vivo identity of tRNA discriminator mutants, *J. Mol. Biol.* 228 (1992) 1055–1062.
- [18] H. Himeno, T. Hasegawa, T. Ueda, K. Watanabe, M. Shimizu, Conversion of aminoacylation specificity from tRNA<sup>Tyr</sup> to tRNA<sup>Ser</sup> in vitro, *Nucleic Acids Res.* 18 (1990) 6815–6819.
- [19] C.L. Quinn, N. Tao, P. Schimmel, Species-specific microhelix aminoacylation by a eukaryotic pathogen tRNA synthetase dependent on a single base pair, *Biochem.* 34 (1995) 12489–12495.
- [20] L. Bare, O.C. Uhlenbeck, Aminoacylation of anticodon loop substituted yeast tyrosine transfer RNA, *Biochem.* 24 (1985) 2354–2360.



- [21] L.A. Bare, O.C. Uhlenbeck, Specific substitution into the anticodon loop of yeast tyrosine transfer RNA, *Biochem. 25* (1986) 5825–5830.
- [22] C.P. Lee, U.L. RajBhandary, Mutants of *Escherichia coli* initiator tRNA which suppress amber codons in *Saccharomyces cerevisiae* and are aminoacylated with tyrosine by yeast extracts, *Proc. Natl. Acad. Sci. USA* 88 (1991) 11378–11382.
- [23] P. Fechter, J. Rudinger-Thirion, A. Théobald-Dietrich, R. Giegé, Identity of tRNA for yeast tyrosyl-tRNA synthetase: tyrosylation is more sensitive to identity nucleotides than to structural features, *Biochem. 39* (2000) 1725–1733.
- [24] P. Fechter, J. Rudinger-Thirion, M. Tukalo, R. Giegé, Major tyrosine identity determinants in *Methanococcus jannaschii* and *Saccharomyces cerevisiae* tRNA<sup>Tyr</sup> are conserved but expressed differently, *Eur. J. Biochem. 268* (2001) 761–767.
- [25] B.A. Steer, P. Schimmel, Major anticodon-binding region missing from an archaeobacterial tRNA synthetase, *J. Biol. Chem. 274* (1999) 35601–35606.
- [26] T.A. Kleeman, D. Wei, K.L. Simpson, E.A. First, Human tyrosyl-tRNA synthetase shares amino acid sequence homology with a putative cytokine, *J. Biol. Chem. 272* (1997) 14420–14425.
- [27] K. Wakasugi, C.L. Quinn, N. Tao, P. Schimmel, Genetic code in evolution: switching species-specific aminoacylation with a peptide transplant, *EMBO J. 17* (1998) 297–305.
- [28] M. Helm, H. Brulé, D. Friede, R. Giegé, J. Pütz, C. Florentz, Search for characteristic structural features of mammalian mitochondrial tRNAs, *RNA* 6 (2000) 1356–1379.
- [29] L. Bonnefond, M. Frugier, R. Giegé, J. Rudinger-Thirion, Human mitochondrial TyrRS disobeys the tyrosine identity rules, *RNA* 8 (11) (2005) (in press).
- [30] R. Giegé, M. Frugier, J. Rudinger, tRNA mimics, *Curr. Opin. Struct. Biol. 8* (1998) 286–293.
- [31] P. Fechter, J. Rudinger-Thirion, C. Florentz, R. Giegé, Novel features in the tRNA-like world of plant viral RNAs, *Cell. Mol. Life Sci. 58* (2001) 1547–1561.
- [32] T.C. Hall, D.S. Shih, P. Kaesberg, Enzyme-mediated binding of tyrosine to brome-mosaic-virus ribonucleic acid, *Biochem. J. 129* (1972) 969–976.
- [33] R.L. Joshi, S. Joshi, F. Chapeville, A.L. Haenni, tRNA-like structures of plant viral RNAs: conformational requirements for adenylation and aminoacylation, *EMBO J. 2* (1983) 1123–1127.
- [34] V. Perret, C. Florentz, T. Dreher, R. Giegé, Structural analogies between the 3' tRNA-like structure of brome mosaic virus RNA and yeast tRNA<sup>Tyr</sup> revealed by protection studies with yeast tyrosyl-tRNA synthetase, *Eur. J. Biochem. 185* (1989) 331–339.
- [35] P. Fechter, R. Giegé, J. Rudinger-Thirion, Specific tyrosylation of the bulky tRNA-like structure of Brome Mosaic Virus RNA relies solely on identity nucleotides present in its amino acid accepting domain, *J. Mol. Biol. 309* (2001) 387–399.
- [36] R.A. Akins, A.M. Lambowitz, A protein required for splicing group I introns in *Neurospora* mitochondria is mitochondrial tyrosyl-tRNA synthetase or a derivative thereof, *Cell* 50 (1987) 331–345.
- [37] M.G. Caprara, G. Mohr, A.M. Lambowitz, A tyrosyl-tRNA synthetase protein induces tertiary folding of the group I intron catalytic core, *J. Mol. Biol. 257* (1996) 512–531.
- [38] P.J. Paukstelis, R. Coon, L. Madabusi, J. Nowakowski, A. Monzingo, J. Robertus, A.M. Lambowitz, A tyrosyl-tRNA synthetase adapted to function in group I intron splicing by acquiring a new RNA binding surface, *Mol. Cell* 17 (2005) 417–428.
- [39] M.G. Caprara, V. Lehnert, A.M. Lambowitz, E. Westhof, A tyrosyl-tRNA synthetase recognizes a conserved tRNA-like structural motif in the group I intron catalytic core, *Cell* 87 (1996) 1135–1145.
- [40] J.I. Guijarro, A. Pintar, A. Prochnicka-Chalufour, V. Guez, B. Gilquin, H. Bedouelle, M. Delepierre, Structure and dynamics of the anticodon arm binding domain of *Bacillus stearothermophilus* tyrosyl-tRNA synthetase, *Structure* 10 (2002) 311–317.
- [41] O. Salazar, B. Sagredo, E. Jedlicki, D. Söll, I. Weygand-Durasevic, O. Orellana, *Thiobacillus ferrooxidans* tyrosyl-tRNA synthetase functions in vivo in *Escherichia coli*, *J. Bacteriol. 176* (1994) 4409–4415.
- [42] Y.I. Wolf, L. Aravind, N.V. Grishin, E.V. Koonin, Evolution of aminoacyl-tRNA synthetases-analysis of unique domain architectures and phylogenetic trees reveals a complex history of horizontal gene transfer events, *Genome Res. 9* (1999) 689–710.
- [43] M.W. Gray, Origin and evolution of mitochondrial DNA, *Annu. Rev. Cell Biol. 5* (1989) 25–50.
- [44] L. Bonnefond, A. Fender, J. Rudinger-Thirion, R. Giegé, C. Florentz, M. Sissler, Toward the full set of human mitochondrial aminoacyl-tRNA synthetases: characterization of AspRS and TyrRS, *Biochemistry* 44 (2005) 4805–4816.
- [45] T.N. Bhat, D.M. Blow, P. Brick, J. Nyborg, Tyrosyl-tRNA synthetase forms a mononucleotide-binding fold, *J. Mol. Biol. 158* (1982) 699–709.
- [46] P. Brick, D.M. Blow, Crystal structure of a deletion mutant of a tyrosyl-tRNA synthetase complexed with tyrosine, *J. Mol. Biol. 194* (1987) 287–297.
- [47] P. Brick, T.N. Bhat, D.M. Blow, Structure of tyrosyl-tRNA synthetase refined at 2.3 Å resolution. Interaction of the enzyme with tyrosyl adenylate intermediate, *J. Mol. Biol. 208* (1989) 83–98.
- [48] X. Qiu, C.A. Janson, W.W. Smith, S.M. Green, P. McDevitt, K. Johanson, P. Carter, M. Hibbs, C. Lewis, A. Chalker, A. Fosberry, J. Lalonde, J. Berge, P. Brown, C.S.V. Houge-Frydrych, R.L. Jarvest, Crystal structure of *Staphylococcus aureus* tyrosyl-tRNA synthetase in complex with a class of potent and specific inhibitors, *Protein Sci. 10* (2001) 2008–2016.
- [49] T. Kobayashi, T. Takimura, R. Sekine, K. Vincent, K. Kamata, K. Sakamoto, S. Nishimura, S. Yokoyama, Structural snapshots of the KMSKS loop rearrangement for amino acid activation by bacterial tyrosyl-tRNA synthetase, *J. Mol. Biol. 346* (2005) 105–117.
- [50] M.A. Markus, R.B. Gerstner, D.E. Draper, D.A. Torchia, The solution structure of ribosomal protein S4 Δ41 reveals two subdomains and a positively charged surface that may interact with RNA, *EMBO J. 17* (1998) 4559–4571.
- [51] L. Lin, S.P. Hale, P. Schimmel, Aminoacylation error correction, *Nature* 384 (1996) 33–34.
- [52] A. Yaremchuk, I. Krikilivyi, M. Tukalo, S. Cusack, Class I tyrosyl-tRNA synthetase has a class II mode of cognate tRNA recognition, *EMBO J. 21* (2002) 3829–3840.
- [53] M.M.Y. Wayne, G. Winter, A.J. Wilkinson, A.R. Fersht, Deletion mutagenesis using an 'M13 splint': the N-terminal structural domain of tyrosyl-tRNA synthetase (*B. stearothermophilus*) catalyses the formation of tyrosyl adenylate, *EMBO J. 2* (1983) 1827–1829.
- [54] P. Carter, H. Bedouelle, G. Winter, Construction of heterodimer tyrosyl-tRNA synthetase shows tRNA<sup>Tyr</sup> interacts with both subunits, *Proc. Natl. Acad. Sci. USA* 83 (1986) 1189–1192.
- [55] H. Bedouelle, G. Winter, A model of synthetase/transfer RNA interaction as deduced by protein engineering, *Nature* 320 (1986) 371–373.
- [56] E. Labouze, H. Bedouelle, Structural and kinetic bases for the recognition of tRNA<sup>Tyr</sup> by tyrosyl-tRNA synthetase, *J. Mol. Biol. 205* (1989) 729–735.
- [57] Y. Xin, W. Li, D.S. Dwyer, E.A. First, Correlating amino acid conservation with function in tyrosyl-tRNA synthetase, *J. Mol. Biol. 303* (2000) 287–298.
- [58] X.L. Yang, R.J. Skene, D.E. McRee, P. Schimmel, Crystal structure of a human aminoacyl-tRNA synthetase cytokine, *Proc. Natl. Acad. Sci. USA* 99 (2002) 15369–15374.
- [59] V.A. Ilyin, B. Temple, M. Hu, G. Li, Y. Yin, P. Vachette, C.W. Carter Jr., 2.9 Å crystal structure of ligand-free tryptophanyl-tRNA synthetase: domain movements fragment the adenine nucleotide binding site, *Protein Sci. 9* (2000) 218–231.

- [60] T. Kobayashi, O. Nureki, R. Ishitani, A. Yaremchuk, M. Tukalo, S. Cusack, K. Sakamoto, S. Yokoyama, Structural basis for orthogonal tRNA specificities of tyrosyl-tRNA synthetases for genetic code expansion, *Nat. Struct. Biol.* 10 (2003) 425–432.
- [61] S. Nair, L. Ribas de Pouplana, F. Houman, A. Avruch, X. Shen, P. Schimmel, Species-specific tRNA recognition in relation to tRNA synthetase contact residues, *J. Mol. Biol.* 269 (1997) 1–9.
- [62] R. Giegé, Genetic code expansion, *Nat. Struct. Biol.* 10 (2003) 414–416.
- [63] H. Edwards, P. Schimmel, An *E. coli* aminoacyl-tRNA synthetase can substitute for yeast mitochondrial enzyme function in vivo, *Cell* 51 (1987) 643–649.
- [64] U. Kamper, U. Kück, A.D. Cherniack, A.M. Lambowitz, The mitochondrial tyrosyl-tRNA synthetase of *Podospira anserina* is a bifunctional enzyme active in protein synthesis and RNA splicing, *Mol. Cell. Biol.* 12 (1992) 499–511.
- [65] S. Cusack, M. Härtlein, R. Leberman, Sequence, structural and evolutionary relationships between class 2 aminoacyl-tRNA synthetases, *Nucleic Acids Res.* 19 (1991) 3489–3498.
- [66] P. Schimmel, L. Ribas de Pouplana, RNA transfer, from minihelix to genetic code, *Cell* 81 (1995) 983–986.
- [67] L. Ribas de Pouplana, M. Frugier, C.L. Quinn, P. Schimmel, Evidence that two present-day components needed for the genetic code appeared after nucleated cells separated from eubacteria, *Proc. Natl. Acad. Sci. USA* 93 (1996) 166–170.
- [68] J.R. Brown, F.T. Robb, R. Weiss, W.F. Doolittle, Evidence for the early divergence of tryptophanyl- and tyrosyl-tRNA synthetases, *J. Mol. Evol.* 45 (1997) 9–16.
- [69] Y. Diaz-Lazcoz, J.C. Aude, P. Nitschké, H. Chiapello, C. Landès-Devauchelle, J.L. Risler, Evolution of genes, evolution of species: the case of aminoacyl-tRNA synthetases, *Mol. Biol. Evol.* 15 (1998) 1548–1561.
- [70] L. Aravind, E.V. Koonin, Novel predicted RNA-binding domains associated with the translation machinery, *J. Mol. Evol.* 48 (1999) 291–302.
- [71] J. Liu, X.L. Yang, K.L. Ewalt, P. Schimmel, Mutational switching of a yeast tRNA synthetase into a mammalian-like synthetase cytokine, *Biochem.* 41 (2002) 14232–14237.
- [72] P. Glaser, F. Kunst, M. Debarbouille, A. Vertes, A. Danchin, R. Dedonder, A gene encoding a tyrosine tRNA synthetase is located near *sacS* in *Bacillus subtilis*, *DNA Seq.* 1 (1991) 251–261.
- [73] J.C. Salazar, R. Zuniga, C. Lefimil, D. Söll, O. Orellana, Conserved amino acids near the carboxy terminus of bacterial tyrosyl-tRNA synthetase are involved in tRNA and Tyr-AMP binding, *FEBS Lett.* 491 (2001) 257–260.
- [74] D. Raoult, S. Audic, C. Robert, C. Abergel, P. Renesto, H. Ogata, B. La Scola, M. Suzan, J.M. Claverie, The 1.2-mb genome sequence of Mimivirus, *Sci.* 306 (2004) 1344–1350.
- [75] J.C. Robinson, P. Kerjean, M. Mirande, Macromolecular assemblage of aminoacyl-tRNA synthetases: quantitative analysis of protein-protein interactions and mechanism of complex assembly, *J. Mol. Biol.* 304 (2000) 983–994.
- [76] K. Wakasugi, P. Schimmel, Two distinct cytokines released from a human aminoacyl-tRNA synthetase, *Sci.* 284 (1999) 147–150.
- [77] A.D. Cherniack, G. Garriga, J.D. Kittle Jr., R.A. Akins, A.M. Lambowitz, Function of *Neurospora* mitochondrial tyrosyl-tRNA synthetase in RNA splicing requires an idiosyncratic domain not found in other synthetase, *Cell* 62 (1990) 745–755.
- [78] B.H. Berg, Purification of aminoacyl-tRNA synthetase kinase activities associated with threonyl- and tyrosyl-tRNA synthetases isolated from Bom:NMRI mouse liver, *Biochem. Mol. Biol. Int.* 29 (1993) 949–958.
- [79] M.X. Guan, Cytoplasmic tyrosyl-tRNA synthetase rescues the defect in mitochondrial genome maintenance caused by the nuclear mutation *mgm104-1* in the yeast *Saccharomyces cerevisiae*, *Mol. Gen. Genet.* 255 (1997) 525–532.
- [80] T.A. Cropp, P.G. Schultz, An expanding genetic code, *Trends Genet.* 20 (2004) 625–630.
- [81] L. Wang, P.G. Schultz, Expanding the genetic code, *Angew. Chem. Int. Ed. Engl.* 44 (2005) 34–66.
- [82] J. Soutourina, S. Blanquet, P. Plateau, D-tyrosyl-tRNA(Tyr) metabolism in *Saccharomyces cerevisiae*, *J. Biol. Chem.* 275 (2000) 11626–11630.
- [83] L. Wang, P.G. Schultz, A general approach for the generation of orthogonal tRNAs, *Chem. Biol.* 8 (2001) 883–890.
- [84] L. Wang, A. Brock, B. Herderich, P.G. Schultz, Expanding the genetic code of *Escherichia coli*, *Science* 292 (2001) 498–500.
- [85] A.K. Kowal, C. Kohrer, U.L. RajBhandary, Twenty-first aminoacyl-tRNA synthetase-suppressor tRNA pairs for possible use in site-specific incorporation of amino acid analogues into proteins in eukaryotes and in eubacteria, *Proc. Natl. Acad. Sci. USA* 98 (2001) 2268–2273.
- [86] I. Hirao, T. Ohtsuki, T. Fujiwara, T. Mitsui, T. Yokogawa, T. Okuni, H. Nakayama, K. Takio, T. Yabuki, T. Kigawa, K. Kodama, K. Nishikawa, S. Yokoyama, An unnatural base pair for incorporating amino acid analogs into proteins, *Nat. Biotechnol.* 20 (2002) 177–182.
- [87] L. Wang, P.G. Schultz, Expanding the genetic code, *Chem. Commun. (Camb.)* (2002) 1–11.
- [88] T. Kobayashi, K. Sakamoto, T. Takimura, R. Sekine, K. Vincent, K. Kamata, S. Nishimura, S. Yokoyama, Structural basis of nonnatural amino acid recognition by an engineered aminoacyl-tRNA synthetase for genetic code expansion, *Proc. Natl. Acad. Sci. USA* 102 (2005) 1366–1371.
- [89] K. Sakamoto, A. Hayashi, A. Sakamoto, D. Kiga, H. Nakayama, A. Soma, T. Kobayashi, M. Kitabatake, K. Takio, K. Saito, M. Shirouzu, I. Hirao, S. Yokoyama, Site-specific incorporation of an unnatural amino acid into proteins in mammalian cells, *Nucleic Acids Res.* 30 (2002) 4692–4699.
- [90] D. Kiga, K. Sakamoto, K. Kodama, T. Kigawa, T. Matsuda, T. Yabuki, M. Shirouzu, Y. Harada, H. Nakayama, K. Takio, Y. Hasegawa, Y. Endo, I. Hirao, S. Yokoyama, An engineered *Escherichia coli* tyrosyl-tRNA synthetase for site-specific incorporation of an unnatural amino acid into proteins in eukaryotic translation and its application in a wheat germ cell-free system, *Proc. Natl. Acad. Sci. USA* 99 (2002) 9715–9720.
- [91] W.L. DeLano, The PyMOL molecular graphics system, Delano Scientific (<http://www.pymol.org>): San Carlos, CA, USA, 2002.
- [92] O. Poirot, K. Suhre, C. Abergel, E. O’Toole, C. Notredame, 3DCoffee@igs: a web server for combining sequences and structures into multiple sequence alignment, *Nucleic Acids Res.* 32 (2004) W37–W40.
- [93] J. Felsenstein, PHYLIP (Phylogeny Inference Package) version 3.62, distributed by the author, Department of Genetics, University of Washington, Seattle, 1993.



La TyrRS est à la fois la première synthétase dont la structure cristallographique a été déterminée (Brick et Blow, 1987) et aussi celle dont la structure a été résolue chez le plus grand nombre d'organismes, avec plus de 30 structures déposées à la PDB (Protein Data Bank, [www.pdb.org](http://www.pdb.org)). Il s'agit de TyrRS seules ou en complexes avec ses différents substrats (acide aminé, ATP, ARNt) et dans plusieurs cas, de mutants de TyrRS mis au point pour l'introduction d'acides aminés non naturels dans des protéines (**cf. C.4.**). Le **tableau 2** liste les différentes structures de TyrRS connues à ce jour.

Les nombreuses séquences et structures de TyrRS disponibles ont permis d'en dessiner l'arbre phylogénétique lors d'études sur l'évolution des aaRS (Diaz-Lazcoz *et al.*, 1998; Wolf *et al.*, 1999; Woese *et al.*, 2000; O'Donoghue et Luthey-Schulten, 2003; Brindefalk *et al.*, 2007). L'arbre présenté en **figure 10** est le plus récent. L'existence d'au moins deux sous-types de TyrRS eubactériennes a été suggérée après la découverte de la présence de deux TyrRS dans certaines bactéries (Glaser *et al.*, 1991). Dans l'arbre phylogénétique présenté ici, les TyrRS mitochondriales d'animaux et de fongi constituent un groupe distinct parmi les TyrRS eubactériennes et proche du 1<sup>er</sup> groupe de TyrRS eubactériennes comprenant *E. coli* et *B. subtilis*. Les TyrRS cytoplasmiques se regroupent avec les TyrRS d'archaea comme attendu, mais au sein de ce groupe les TyrRS d'animaux et de fongi se distinguent très nettement de celles de plantes. Les TyrRS cytoplasmiques et mitochondriales étant codées par le génome nucléaire, l'observation de la même répartition des espèces pour chaque type de TyrRS souligne la stabilité du génome nucléaire des eucaryotes.

Groupe	PDB	Organisme	A	Groupe	R	R-free	Ligands	Molécule cristallisée	Auteurs	
Bactéries	ZTS1	<i>B. stearothermophilus</i>	2,3	P3121	0,228	n/a	intermédiaire tyrosyl-adénylate		Brick et al., 1989	
	3TS1	<i>B. stearothermophilus</i>	2,7	P3121	0,21	n/a	intermédiaire tyrosyl-adénylate		Brick et al., 1989	
	4TS1	<i>B. stearothermophilus</i>	2,5	P21(P1211)	0,187	n/a	tyrosine		Brick et al., 1987	
	1TYD	<i>B. stearothermophilus</i>	2,5	P3121	0,223	n/a	intermédiaire tyrosyl-adénylate	T51S		
	1TYA	<i>B. stearothermophilus</i>	2,8	P3121	0,205	n/a		T51A		
	1TYB	<i>B. stearothermophilus</i>	2,5	P3121	0,23	n/a		T51G		
	1TYC	<i>B. stearothermophilus</i>	2,5	P3121	0,23	n/a		T51P		
	1JII	<i>S. aureus</i>	3,2	C2221	0,269	0,335	SB-219383		Qiu et al., 2001	
	1JJ	<i>S. aureus</i>	3,2	C2221	0,257	0,305	SB-239629		Qiu et al., 2001	
	1JK	<i>S. aureus</i>	2,8	I212121	0,205	0,305	SB-243545		Qiu et al., 2001	
	1JL	<i>S. aureus</i>	2,2	I212121	0,183	0,221	SB-284485		Qiu et al., 2001	
	1JH3	<i>B. stearothermophilus</i>						domaine C-terminal (321-419)	Guljarro et al., 2002	
	1H3F	<i>T. thermophilus</i>	2	P212121	0,234	0,261	tyrosinol		Yaremchuk et al., 2002	
	1H3E	<i>T. thermophilus</i>	2,9	P3121	0,219	0,269	ARNt, ATP et tyrosinol		Yaremchuk et al., 2002	
1X8X	<i>E. coli</i>	2	P3121	0,194	0,22	tyrosine		Kobayashi et al., 2005a		
1VBM	<i>E. coli</i>	2,7	P3221	0,246	0,284	Tyr-AMS	résidus 5 à 322	Kobayashi et al., 2005a		
1VBN	<i>E. coli</i>	2,7	P3221	0,217	0,283	Tyr-AMS	Y37V, Q195C	Kobayashi et al., 2005b		
1WQ3	<i>E. coli</i>	2	P3121	0,18	0,219	3-iodo-L-tyrosine	Y37V, Q195C	Kobayashi et al., 2005b		
1WQ4	<i>E. coli</i>	2	P3121	0,187	0,224	L-tyrosine	Y37V, Q195C	Kobayashi et al., 2005b		
Archaea	1J1U	<i>M. jannaschii</i>	1,95	P3121	0,188	0,239	ARNt et L-tyrosine	A158D, Q32Y, R107E, P162L	Zhang et al., 2005	
	1U7X	<i>M. jannaschii</i>	3	P212121	0,237	0,308	o-méthyl-L-tyrosine		Zhang et al., 2005	
	1U7D	<i>M. jannaschii</i>	2,65	P212121	0,229	0,277			Zhang et al., 2005	
	1ZH6	<i>M. jannaschii</i>	2,5	P43212	0,213	0,278	p-acétylphénylalanine	Y32L,D158G,I159C,L162R	Turner et al., 2005	
	2AG6	<i>M. jannaschii</i>	1,9	P43212	0,211	0,247	p-bromo-L-phénylalanine	Y32L,E107S,D158P,I159L,L162E	Turner et al., 2006	
	1ZHO	<i>M. jannaschii</i>	1,9	P43212	0,22	0,275	L-3-(2-naphtyl)alanine	Y32L,E107S,D158P,I159A,L162Q,A167V	Turner et al., 2006	
	2CYA	<i>A. permix</i>	2,2	P43212	0,211	0,238			Kuratani et al., 2006	
	2CVB	<i>A. fulgidus</i>	1,8	P21(P1211)	0,202	0,24	L-tyrosine		Kuratani et al., 2006	
	2CYC	<i>P. horikoshii</i>	2,2	P212121	0,171	0,217	L-tyrosine		Kuratani et al., 2006	
	Mimiviridae	215B	<i>Mimivirus</i>	2,2	P212121			ATP et tyrosinol	Abergel et al., 2007 (soumis)	
	Eucaryotes	2DLC	<i>S. cerevisiae</i>	2,4	P41212	0,245	0,289	ARNt et Tyr-AMPN		Tsunoda et al., 2007
		1Y42	<i>N. crassa</i>	1,95	C2(C121)	0,179	0,235		CYT-18 sans domaine C-terminal	Paukstels et al., 2005
		1N3L	<i>H. sapiens</i>	1,18	P21212	0,18	0,223		mini TyrRS	Yang et al., 2002
		1NTG	<i>H. sapiens</i>	2,21	P1	0,251	0,296		domaine EMAP II	Yang et al., 2003
	1Q11	<i>H. sapiens</i>	1,6	P21212	0,197	0,217	tyrosinol	mini TyrRS	Yang et al., 2003	
	2PID	<i>H. sapiens</i>	2,2	P43212			Tyr-AMS	mitochondriale sans domaine S4	Bonnefond et al., 2007 (soumis)	

Tableau 2. Structures cristallographiques de TyrRS connues déposées à la PDB



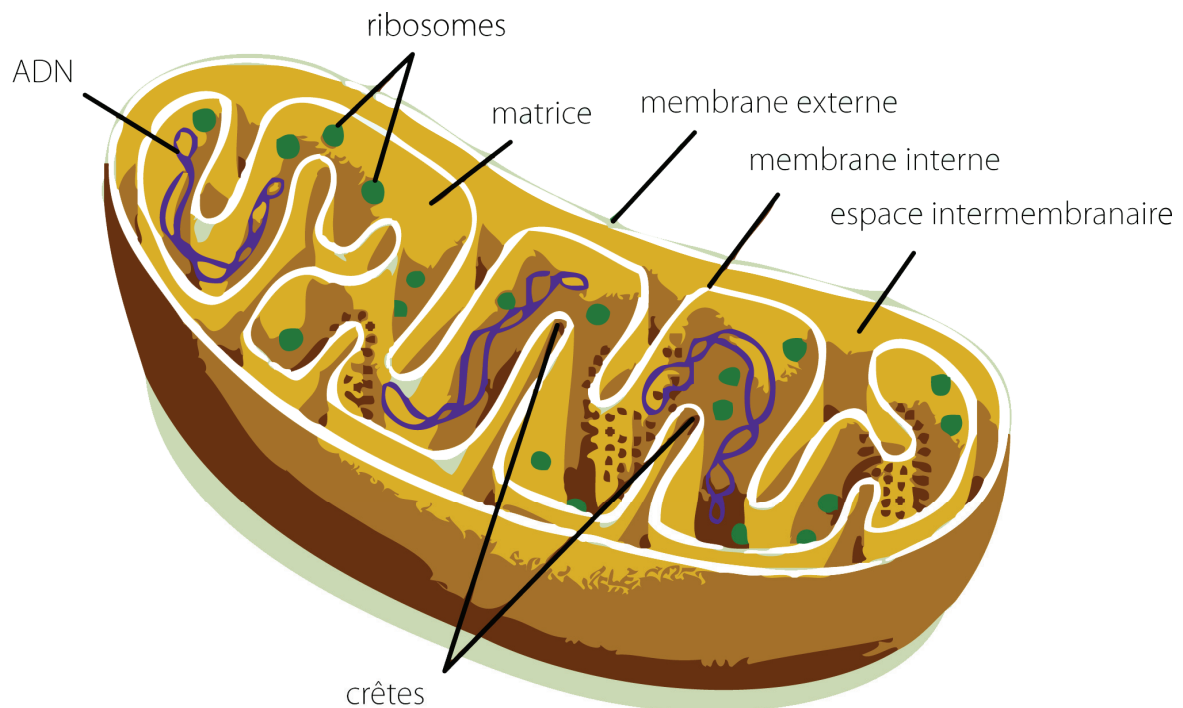
**Figure 10. Arbre phylogénétique des TyrRS**

La construction de cet arbre a été faite à partir d'un lot représentatif d'espèces de bactéries, d'archaea et d'eucaryotes. Les chiffres indiqués sur les branches correspondent aux valeurs de bootstrap supérieures à 75% par les méthodes de ML (Maximum Likelihood) et NJ (Neighbor Joining), respectivement. Les couleurs indiquent des catégories d'espèces : rouge pour les  $\alpha$ -protéobactéries, jaune pour les mitochondries, bleu pour les eucaryotes, vert pour les mitochondries de plantes et les chloroplastes. L'arbre est enraciné en utilisant le même "out-group" que dans des analyses antérieures (Woese et al., 2000). Figure de Brindefalk et al. (2007).

## E. Les aaRS mitochondriales humaines

### 1. La mitochondrie humaine

#### 1.1. Structure et fonction



**Figure 11. Organisation schématique de la mitochondrie humaine et de ses constituants**  
Illustration extraite de l'Encyclopedia Britannica.

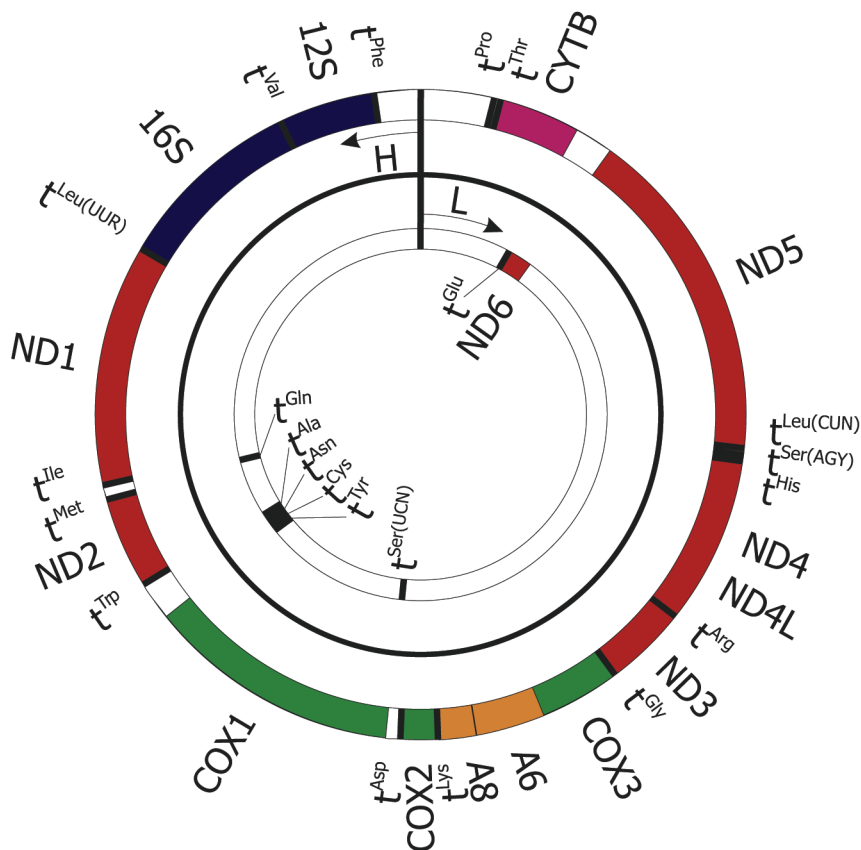
Le mot mitochondrie dérive du grec *mitos*, « filament », et *chondros*, « graine » en raison de l'aspect de cet organite au microscope optique et électronique. La mitochondrie, organelle de quelques dixièmes de micromètre, est présente en nombre variable dans le cytosol. Les mitochondries ne sont pas des organites statiques : elles se scindent ou, au contraire, fusionnent activement ce qui explique leur polymorphisme. Elles sont délimitées par deux enveloppes membranaires. La membrane externe est perméable à toutes les molécules de 5 kDa ou moins grâce à la présence de porines. Elle contient aussi des translocases, transporteurs protéiques, impliquées dans l'import des protéines (Translocase of the Outer Membrane, TOM). La membrane interne est imperméable aux ions et autres molécules, mais contient des translocases (Translocase of the Inner Membrane, TIM) pour l'import des protéines. Entre les deux membranes se trouve un espace inter-membranaire

dont la composition est essentielle au fonctionnement de la mitochondrie. C'est au niveau de la membrane interne qu'a lieu la phosphorylation oxydative qui permet de produire de l'ATP, faisant de la mitochondrie la centrale énergétique de la cellule. La présence de replis formant des crêtes permet d'augmenter la surface de la membrane interne et le volume de l'espace inter-membranaire, ce qui accroît le rendement de production d'ATP. Enfin, la cavité délimitée par la membrane interne correspond à la matrice mitochondriale. Elle contient le génome mitochondrial et sa machinerie traductionnelle, les ARNm et ARNt mitochondriaux et des enzymes dont ceux du cycle de Krebs notamment. En plus de son rôle dans l'approvisionnement de la cellule en ATP, la mitochondrie est impliquée dans d'autres réactions métaboliques : néoglucogenèse, cycle de l'urée, dégradation des acides gras, biosynthèse des hormones stéroïdes. La mitochondrie participe également, avec le réticulum endoplasmique, à la régulation de la concentration intracellulaire de calcium. Enfin, elle intervient dans le mécanisme de mort programmée (apoptose). Dans ce processus, la création d'un pore dans la membrane mitochondriale externe permet la fuite de cytochrome c dans le cytoplasme. Le cytochrome c participe alors à l'activation de protéases (caspases) qui détruisent d'importants composants moléculaires du noyau et du cytoplasme, conduisant à la mort de la cellule.

## **1.2. Génome**

La mitochondrie contient plusieurs copies d'un ADN circulaire (ADNmt) double brin d'une taille de 16 569 bases dont la composition en nucléotides est différente de celle du génome nucléaire. Ce génome est transmis de façon héréditaire par la mère par l'intermédiaire des mitochondries présentes dans l'ovocyte. Les deux brins qui le composent n'ont pas la même teneur en bases. On distingue ainsi un brin dit "lourd", plus riche en G, d'un brin "léger", plus riche en C, U et A. Chaque brin du génome est répliqué à partir de son origine de réplication propre, de manière asymétrique (Clayton, 1982) ou coordonnée (Holt et al., 2000). L'autre particularité de ce génome est son extrême compaction. Il ne contient aucune région intergénique, ni séquence intronique et code pour seulement 13 protéines (uniquement des constituants de la chaîne respiratoire), 2 ARNr et 22 ARNt. La seule région non codante (boucle D) est nécessaire à la réplication et à la transcription du génome. Toutes les autres protéines mitochondriales sont codées par le génome nucléaire de la cellule et importées du cytoplasme. La transcription de l'ADNmt est bidirectionnelle et indépendante pour chaque brin. Elle est effectuée par une ARN polymérase de type  $T_3/T_7$  et génère trois transcrits polycistroniques. Ces transcrits sont maturés par clivage au niveau des

ARNt par les RNase P (en amont) et RNase Z (en aval) (Levinger et al., 2004). Les ARNt libérés sont modifiés et la séquence CCA ajoutée. Les ARNm sont eux polyadénylés, mais ne possèdent ni coiffe en 5', ni séquence particulière en amont du site d'initiation (Attardi et Montoya, 1983).



**Figure 12. Organisation du génome mitochondrial humain**

Les flèches dénommées H et L indiquent le sens de transcription des brins "lourd" et "léger", respectivement. Le cercle extérieur est le transcrite "léger" et le cercle interne le transcrite "lourd". Les ARNr et ARNm sont en couleurs. Les barres noires désignent les ARNt qui ponctuent le génome mitochondrial. Figure de Levinger et al. (2004).

### 1.3. Synthèse protéique

La traduction des ARNm mitochondriaux est étroitement associée à la membrane mitochondriale interne (Liu et Spremulli, 2000) où se localisent les ribosomes mitochondriaux. Son mécanisme est proche des systèmes eubactériens mais présente certaines particularités. Le code génétique mitochondrial tout d'abord présente des variations par rapport au code "universel" qui diffèrent suivant les organismes. Dans les génomes des mitochondries de mammifères, le codon stop opale UGA est réassigné en codon tryptophane, les codons isoleucine AUA et AUU en codons méthionine et les codons arginine

AGA et AGG en codons stop. Étant donné le nombre restreint d'ARNt disponibles pour décoder la soixantaine de codons existants (un par acide aminé excepté pour la leucine et sérine), une interaction codon-anticodon limitée à deux positions sur les trois pourrait permettre une plus grande flexibilité de reconnaissance ARNm/ARNt (revue par Florentz et al., 2003). Les ribosomes mitochondriaux se distinguent par leur faible contenu en ARNr (33% contre 66% dans le ribosome bactérien et 60% dans le ribosome cytoplasmique).

#### **1.4. Origine endosymbiotique**

Une  $\alpha$ -protéobactérie serait l'endosymbiote original, ancêtre de la mitochondrie actuelle (Margulis, 1981). Il aurait subi au cours de l'évolution un transfert massif de ses gènes vers le génome nucléaire de l'hôte pour réduire le génome mitochondrial de façon extrême. Pourtant l'analyse des protéomes mitochondriaux révèle seulement 20% de protéines d'origine  $\alpha$ -protéobactérienne, contre 30 % d'origine bactérienne autre et 50 % de protéines ne présentant pas d'homologues bactériens. Parallèlement, l'étude phylogénétique des aaRS suggère que le génome nucléaire ancestral contenait déjà des gènes de type eubactérien auquel seul un jeu limité de gènes de l' $\alpha$ -protéobactérie endocytée (principalement des composants de la chaîne respiratoire) a été ajouté et sélectivement maintenu (Brindefalk et al., 2007). Leur maintien dans le génome nucléaire serait lié au caractère très hydrophobe des protéines correspondantes, empêchant leur import à travers la membrane mitochondriale. Ainsi, les événements d'endocytose de bactéries et de transfert de gènes semblent avoir été nombreux au cours de l'évolution des cellules eucaryotes comme le montre la grande variété de génomes mitochondriaux existants.

## **2. Les systèmes d'aminocyclation mitochondriaux humains**

Dans la mitochondrie, les partenaires des systèmes d'aminocyclation ont des origines génétiques différentes, les ARNt étant codés par le génome mitochondrial alors que les synthétases sont codées par le génome nucléaire de la cellule puis importées.

## **2.1. Les ARNt mitochondriaux**

### **2.1.1. Originalités de séquence**

Le génome mitochondrial humain code pour 22 ARNt soit un par acide aminé, à l'exception de la leucine et de la sérine qui sont codées par 2 ARNt (Anderson et al., 1981). Ce lot d'ARNt est réduit mais suffisant pour assurer la synthèse protéique mitochondriale, et ne nécessite pas l'import d'ARNt additionnel du cytoplasme comme il a été observé pour l'ARNt<sup>Lys</sup> chez différentes espèces de marsupiaux (Dörner et al., 2001), levures, plantes et protozoaires (Bhattacharyya et Adhya, 2004). Les gènes des ARNt sont répartis sur les deux brins du génome mitochondrial avec 14 ARNt sur le brin "lourd" et les 8 autres sur le brin "léger" dont la composition en bases est différente. Le brin "lourd" est plus riche en G que le brin "léger". Les ARNt "lourds", codés par le brin "léger", ont une plus grande proportion de paires de bases G•U. Les ARNt "légers" ont peu de G dans les boucles D et T et ont une forte proportion de dinucléotides CpA et UpA qui les rend plus sensibles à la dégradation. De plus, du fait de l'extrême compaction du génome mitochondrial, ces ARNt sont globalement plus petits que leurs homologues cytoplasmiques (de 62 à 78 nucléotides contre 75 à 95).

### **2.1.2. Maturation**

La maturation des ARNt mitochondriaux fait intervenir les RNases P et Z pour la maturation en 5' et 3'. Les ARNt ainsi libérés sont modifiés et le triplet CCA est ajouté par la CCAsase (Levinger et al., 2004). A l'heure actuelle une centaine d'ARNt mitochondriaux (la moitié provenant d'animaux) ont été séquencés et certaines modifications post-transcriptionnelles identifiées (Sprinzl et Vassilenko, 2005). Chez l'homme, la séquence de 6 ARNt mitochondriaux est connue (Lys, Ile, Leu(UUR), Ser(UCN), Ser(AGY) et Pro). Leur analyse a montré qu'ils présentent moins de modifications que leurs homologues cytoplasmiques. La plupart de ces modifications existent chez les ARNt cytoplasmiques d'autres organismes, mais certaines sont présentes à des positions inhabituelles comme la méthylation de l'adénine en position 9 identifiée pour 13 spécificités d'ARNt mitochondriaux. La modification de cette position permet à l'ARNt<sup>Lys</sup> de se replier correctement en "feuille de trèfle" en empêchant la formation d'une paire de bases Watson-Crick (Helm et al., 1999). Au vu du faible nombre d'ARNt mitochondriaux humains séquencés à ce jour, il reste certainement encore beaucoup de modifications à découvrir, à l'image de la découverte récente des uridines modifiées contenant de la taurine dans l'anticodon des ARNt<sup>Lys</sup> et ARNt<sup>Leu</sup> (Suzuki et al., 2002).

### **2.1.3. Structures "bizarres"**

Les ARNt mitochondriaux se distinguent également au niveau structural, conséquence logique de leurs séquences atypiques qui ont leur ont valu le qualificatif de "bizarres"



(Steinberg et Cedergren, 1994). L'exemple le plus extrême est celui de l'ARNt<sup>Ser</sup><sub>AGY</sub> mitochondrial bovin qui ne possède pas de bras D. Généralement, les ARNt mitochondriaux peuvent être repliés en "feuille de trèfle" caractéristique, mais avec d'importantes variations dans la taille de la boucle D et surtout de la boucle T. La région variable des ARNt mitochondriaux est systématiquement restreinte à 4-5 nucléotides, même pour les ARNt dont les homologues eubactériens présentent de grandes régions variables (ARNt<sup>Tyr</sup>, ARNt<sup>Leu</sup> et ARNt<sup>Ser</sup>). Leur structure secondaire présente un nombre plus important de paires de bases A-U et G-U, et très peu de paires de bases G-C, ce qui augmente leur flexibilité. Leur repliement tridimensionnel est à l'heure actuelle encore mal compris. En effet les nucléotides connus et conservés dans les ARNt "classiques" impliqués dans la formation du réseau d'interactions tertiaires sont souvent absents dans les ARNt mitochondriaux. La possibilité d'un réseau d'interaction alternatif propre n'a pas encore pu être rationalisée (Hayashi et al., 1998).

#### **2.1.4. Implications dans des pathologies**

Le taux de mutation extrêmement élevé du génome mitochondrial n'est pas sans conséquence sur le métabolisme cellulaire. Alors que de nombreuses mutations dites polymorphiques sont sans effets sur le fonctionnement de la mitochondrie et contribuent à la diversité des populations (Ingman et al., 2000), beaucoup induisent des pathologies (DiMauro et Moraes, 1993 ; DiMauro et Schon, 2001). A l'heure actuelle, la banque MITOMAP (Kogelnik et al., 1998) recense plus d'une centaine de pathologies corrélées à des mutations dans l'ADN mitochondrial, dont 38% touchent les gènes de protéines, 4% les gènes d'ARNr et 58% les gènes ARNt. Les mutations pathogéniques n'affectent pas tous les ARNt mitochondriaux dans les mêmes proportions. Par contre, il n'y pas de préférence dans le domaine de l'ARNt affecté, les mutations étant réparties de façon homogène sur la quasi-totalité des positions des ARNt (Florentz et al., 2003). Ces mutations peuvent affecter différentes étapes, allant de la biogenèse des ARNt à leur interaction avec leurs différents partenaires protéiques. Ainsi, des mutations affectant la modification post-transcriptionnelle des ARNt (Brulé et al., 1998), leur stabilité (Hao et Moraes, 1997), leur maturation (Levinger et al., 2001) et leur capacité d'aminocyclation ont été mises en évidence. Il est intéressant de noter que les mutations identifiées et analysées jusqu'à présent comme affectant la capacité d'aminocyclation des ARNt mitochondriaux induisent également une déstabilisation de leur structure (Levinger et al., 2004). Le type de pathologie associée est assez vaste et comprend toute une gamme d'encéphalomyopathies, de myopathies mitochondriales, de cytopathies, de surdités et d'affections musculaires (MITOMAP, [www.mitomap.org](http://www.mitomap.org)). La proportion variable de mitochondries affectées par une mutation donnée dans une même

cellule, un type de tissu ou un organe, multiplie les possibilités de pathologies et rend les tentatives de diagnostic périlleuses. En effet, un seuil de mitochondries affectées par cellule est nécessaire pour induire des perturbations métaboliques, on parle d'hétéroplasmie.

## 2.2. Les aminoacyl-ARNt synthétases mitochondriales

Alors que la plupart des activités d'aminacylation mitochondriales ont été détectées (Lynch et Attardi, 1976; Brulé et al., 1998), les gènes des synthétases mitochondriales ne sont étudiés que depuis une dizaine d'années. L'identification des gènes des GlyRS et IleRS cytoplasmiques humaines (Shiba et al., 1994a ; Shiba et al., 1994b) a indirectement permis de découvrir les premières synthétases mitochondriales. Dans le cas de la GlyRS, un même gène code pour la version cytoplasmique et mitochondriale de la synthétase qui ne diffèrent que par la présence d'un signal d'adressage vers la mitochondrie pour la deuxième forme. Ce signal d'adressage se présente sous la forme d'une hélice  $\alpha$  amphiphile riche en résidus chargés et située à l'extrémité N-terminale de la protéine concernée (Hammen et Weiner, 1998). Le même raisonnement a permis d'identifier l'HisRS mitochondriale (O'Hanlon et al., 1995) qui partage le même promoteur bidirectionnel et 70 % d'identité avec son homologue cytoplasmique. Par la suite, huit autres synthétases mitochondriales ont été identifiées et caractérisées, il s'agit de la PheRS (Bullard et al., 1999), LeuRS (Bullard et al., 2000), LysRS (Tolkunova et al., 2000), TrpRS (Jørgensen et al., 2000) et MetRS (Spencer et al., 2004), AspRS et TyrRS (**article n°2**) et AlaRS (Dickey et al., 2006). Ainsi soit l'aaRS mitochondriale est codée par un gène différent de l'aaRS cytoplasmique, c'est le cas pour l'HisRS, PheRS, LeuRS, IleRS, TrpRS, MetRS, AspRS et TyrRS ; soit le même gène code pour les deux aaRS à la fois, c'est le cas pour la GlyRS et LysRS. Pour la GlyRS, deux sites d'initiation de la traduction permettent de produire les deux formes de l'enzyme, alors que dans le cas la LysRS, c'est un épissage alternatif qui entre en jeu (Tolkunova et al., 2000). Au-delà de cette classification binaire, l'analyse des relations phylogénétiques entre les synthétases cytoplasmiques et mitochondriales est beaucoup plus ardue. Prenons l'exemple d'un gène codant à la fois les formes cytoplasmique et mitochondriale d'une synthétase donnée. Ce gène peut être soit le gène de l'endosymbiote, donc d'origine eubactérienne, soit le gène eucaryotique de la synthétase cytoplasmique. Mais le problème se pose alors en amont quant à l'origine du gène eucaryotique, qui peut avoir été remplacé au cours de l'évolution par un gène eubactérien. Parallèlement le gène mitochondrial peut avoir subi plusieurs épisodes de transfert de gène et être plus proche du gène nucléaire. Les scénarios

envisageables sont nombreux et nécessitent des études phylogénétiques poussées pour comprendre l'histoire évolutive des synthétases mitochondriales (Brindefalk et al., 2007).

	taille en aa		oligométrie	classe	référence
	MTS	aaRS mature			
<b>IleRS</b>	?	993	$\alpha$ ?	Ia	(Shiba et al., 1994b)
<b>LeuRS</b>	39	903	$\alpha_2$	Ia	(Bullard et al., 2000)
<b>MetRS</b>	18	575	$\alpha$	Ia	(Spencer et al., 2004)
<b>TrpRS</b>	18	360	$\alpha_2$	Ic	(Jørgensen et al., 2000)
<b>TyrRS</b>	32	435	$\alpha_2$	Ic	<b>Article n°2</b>
<b>SerRS</b>	20	498	$\alpha_2$	IIa	(Yokogawa et al., 2000)
<b>GlyRS</b>	54	685	$\alpha_2$	IIa	(Shiba et al., 1994a; Mudge et al., 1998)
<b>HisRS</b>	?	506	$\alpha_2$ ?	IIa	(O'Hanlon et al., 1995)
<b>AspRS</b>	47	598	$\alpha_2$	IIb	<b>Article n°2</b>
<b>LysRS</b>	16	625	$\alpha_2$	IIb	(Tolkunova et al., 2000)
<b>PheRS</b>	37	451	$\alpha$	IIc	(Bullard et al., 1999)

**Tableau 3. Caractéristiques des aminoacyl-ARNt synthétases mitochondriales humaines**  
L'oligométrie des HisRS et IleRS est prédictive et indiquée à partir de celle des enzymes connues.  
MTS, signal d'adressage à la mitochondrie.

Les aminoacyl-ARNt synthétases mitochondriales partagent les mêmes tailles et oligométries que les aaRS connues, sauf pour la GlyRS mitochondriale qui est homodimérique ( $\alpha_2$ ) et la PheRS mitochondriale qui est monomérique ( $\alpha$ ) et non hétérotétramérique ( $\alpha_2\beta_2$ ) (**tableau 3**). Elles conservent néanmoins toutes les motifs caractéristiques les définissant comme des synthétases de classe I ou II. A ce jour, la seule structure résolue d'une synthétase mitochondriale de mammifère est celle de la SerRS mitochondriale de bœuf (Chimnaronk et al., 2005). Cette synthétase possède deux extensions caractéristiques qui lui permettent de discriminer entre les deux ARNt<sup>Ser</sup> mitochondriaux (avec et sans bras D) présents dans la mitochondrie en adoptant des conformations alternatives. D'un point de vue fonctionnel, ces synthétases présentent des activités catalytiques réduites par rapport à leurs homologues cytoplasmiques ou eubactériennes (**tableau 4**). Pour la PheRS, TrpRS et LeuRS l'efficacité catalytique ( $k_{cat}/K_M$ ) pour l'acide aminé est 100 fois plus faible que pour les enzymes d'*E. coli*, et 250 fois plus faible pour l'ATP (Bullard et al., 1999; Bullard et al., 2000; Jørgensen et al., 2000). Sur des extraits d'ARNt total d'*E. coli*, la LeuRS a une activité spécifique 250 à 400 fois plus faible que ses homologues connus, et la PheRS 20 à 30 fois plus faible que d'autres synthétases de classe II (Bullard et al., 1999). Ces activités réduites pourraient refléter une synthèse protéique peu active dans la mitochondrie ou bien suggérer

l'intervention d'un facteur additionnel qui améliorerait l'aminocyclation en structurant l'ARNt mitochondrial par exemple. Une autre particularité de ces synthétases mitochondriales est leur large spectre d'action. Ces synthétases sont capables d'aminocycler, en plus de leurs propres ARNt, les ARNt cytoplasmiques et eubactériens homologues, alors que ni les synthétases cytoplasmiques ni les eubactériennes ne reconnaissent les ARNt mitochondriaux (Buck et Nass, 1969 ; Lynch et Attardi, 1976; Kumazawa et al., 1991). Cette unilatéralité d'aminocyclation reflète la singularité des ARNt mitochondriaux qui ne sont pas reconnus par des synthétases d'une autre espèce.

<b>Enzyme</b>	<i>substrat</i>	$K_M$ ( $\mu\text{M}$ )	$k_{cat}$ ( $\text{s}^{-1}$ )	$k_{cat}/K_M$ ( $\text{s}^{-1}\cdot\mu\text{M}^{-1}$ )	référence
<b>IleRS</b>	tRNA <sup>Ile</sup> natif	0,127	392 #	3087 #	(Degoul et al., 1998)
	transcrit	2,92	187 #	187 #	(Degoul et al., 1998)
	transcrit	4,3	0,13	0,03	(Kelley et al., 2000)
<b>LeuRS</b>	<i>E. coli</i> natif	14	0,12	0,008	(Bullard et al., 2000)
	transcrit			0,055	(Sohm et al., 2003)
	transcrit			0,003	(Sohm et al., 2003)
	transcrit	17,9	0,024	0,001	(Sohm et al., 2003)
	transcrit UUR	2,3	0,02	0,009	(Wittenhagen et Kelley, 2003)
<b>MetRS</b>	<i>E. coli</i> natif	2,1	11	5,2	(Spencer et al., 2004)
	<i>B. taurus</i> natif	0,15	0,019	0,13	(Spencer et al., 2004)
	transcrit <i>B. taurus</i>	0,079	0,018	0,23	(Spencer et al., 2004)
	transcrit <i>H. sapiens</i>	0,16	0,021	0,13	(Spencer et al., 2004)
<b>TyrRS</b>	<i>E. coli</i>	0,52	0,21	0,4	<b>Article n°2</b>
	transcrit	4,8	0,046	0,01	<b>Article n°2</b>
<b>SerRS</b>	tRNA <sub>GCU</sub> <sup>Ser</sup>	0,37	0,35	0,95	(Yokogawa et al., 2000)
	tRNA <sub>UGA</sub> <sup>Ser</sup>	0,22	0,63	2,86	(Yokogawa et al., 2000)
	tRNA <sub>UGA</sub> <sup>Ser</sup>	0,29	0,67	2,31	(Shimada et al., 2001)
	transcrit <sub>UGA</sub> <sup>Ser</sup>	1,2	0,25	0,21	(Shimada et al., 2001)
<b>LysRS</b>	tRNA <sup>Lys</sup>	0,71	0,003 *		(Yasukawa et al., 2001)
<b>AspRS</b>	<i>E. coli</i>	0,15	0,133	0,89	Article n°2
	transcrit	0,13	0,01	0,08	Article n°2
	transcrit (15°C)	0,19	0,005	0,03	(Fender et al., 2006)
	transcrit	0,72	0,027		(Scheper et al., 2007)
<b>PheRS</b>	<i>E. coli</i>	18	0,11	0,006	(Bullard et al., 1999)

**Tableau 4. Paramètres cinétiques de l'aminocyclation des aaRS mitochondriales humaines**

#  $k_{cat}$  et  $k_{cat}/K_M$  remplacé par  $V_{max}$  (en  $\text{pmol}\cdot\text{min}^{-1}\cdot\text{mg}^{-1}$ ) et  $V_{max}/K_M$  (en  $\text{pmol}\cdot\text{min}^{-1}\cdot\text{mg}^{-1}\cdot\mu\text{M}^{-1}$ ).

\*  $k_{cat}$  remplacé par  $V_{max}$  (en  $\mu\text{M}\cdot\text{s}^{-1}$ )

### 2.3. Perte de spécificité des systèmes d'aminacylation mitochondriaux

Les quelques aaRS mitochondriales humaines étudiées à ce jour ont montré une spécificité relâchée, que ce soit au niveau de la reconnaissance de l'acide aminé ou de l'ARNt.

La discrimination de l'acide aminé est essentielle dans le cas de LeuRS ; son activité d'édition est d'ailleurs vitale pour la survie des bactéries. Pourtant dans la mitochondrie humaine l'activité d'édition est totalement absente sans qu'il y ait pour autant des erreurs lors de la synthèse protéique. En fait, cette perte d'activité est ici compensée par une spécificité accrue pour la leucine (Lue et Kelley, 2005). De la même manière, la PheRS et la ThrRS mitochondriale humaine sont toutes deux dépourvues de leurs domaines d'édition (Dock-Bregeon et al., 2000; Beebe et al., 2004; Korencic et al., 2004; Roy et al., 2005).

Parallèlement l'AspRS, l'AlaRS et la TyrRS mitochondriale humaine présentent une spécificité relâchée pour leurs ARNt homologues. L'AspRS mitochondriale humaine a perdu la spécificité de reconnaissance de la base discriminatrice G73, alors que cet élément est un déterminant majeur dans tous les autres systèmes Asp caractérisés à ce jour. Cette perte résulte de la mutation d'un résidu essentiel pour la reconnaissance de la base G73 dans le site catalytique de l'AspRS (Fender et al., 2006). L'AlaRS mitochondriale humaine, quant à elle, ne reconnaît pas les positions d'identité de l'ARNt<sup>Ala</sup> caractérisées dans les autres organismes (paire de base G3:U70 et base discriminatrice 73) mais mettrait à profit son extrémité C-terminale étendue pour créer un second jeu de contacts avec l'ARNt (Dickey et al., 2006). Le cas de la TyrRS sera développé dans la partie **Résultats**.

Cette dégénérescence de la spécificité de la réaction d'aminacylation témoigne de l'évolution réductrice qui a lieu dans la mitochondrie (Ribas de Pouplana, 2005). La question se pose alors de savoir (i) si la mitochondrie a recours à d'autres mécanismes pour assurer la fidélité de sa synthèse protéique ou (ii) si elle tolère un taux d'erreurs supérieur à ceux des bactéries ou du cytoplasme des eucaryotes ou encore (iii) si le confinement de la traduction mitochondriale et le nombre restreint de ses ARNt ne nécessite pas une spécificité forte. Pour répondre à la première hypothèse, on peut imaginer que la fidélité de la synthèse protéique mitochondriale est atteinte par le contrôle des concentrations en acides aminés dans la matrice mitochondriale. L'intervention d'un facteur d'édition agissant en *trans* (comme les DTD ou AlaX, cf. C.1.2.) est aussi envisageable, bien qu'aucune activité n'ait été détectée dans la mitochondrie humaine jusqu'à présent. La seconde hypothèse d'une fidélité de traduction réduite est aussi plausible, sachant que seulement 13 protéines mitochondriales sur les milliers présentes dans la mitochondrie sont codées par le génome

mitochondrial. Si les protéines incorrectes et mal repliées sont reconnues et prises en charge par le facteur EF-Tu mitochondrial (Suzuki et al., 2007) couplé à un système de dégradation rapide des protéines mitochondriales (Augustin et al., 2005), leur impact sur la mitochondrie peut être limité.

## F. Objectifs de la thèse

Le choix du système d'aminacylation spécifique de la tyrosine dans la mitochondrie humaine n'est pas anodin. Les systèmes d'aminacylation spécifiques de la tyrosine sont étudiés depuis plus de 50 ans dans des bactéries, des archaea et des eucaryotes. Les études fonctionnelles et structurales ont permis de mettre au jour la barrière d'espèce de ce système, mais aussi les principes généraux de la réaction d'aminacylation des ARNt<sup>Tyr</sup> par opposition aux autres systèmes. Il est clairement établi que la présence d'une barrière d'espèce entre systèmes eubactériens et archaea/eucaryotes résulte de la présence de différences au niveau de l'ARNt et de la synthétase. Pourtant la mise en place de cette barrière d'un point de vue évolutif et sa finalité ne sont pas encore comprises. Le système mitochondrial humain possède des caractéristiques à la fois eubactériennes et eucaryotes tant au niveau de son ARNt que de sa synthétase, conséquence de la scission de son génome et de l'évolution séparée de ses constituants cellulaires. Il constitue ainsi un système "hybride", modèle de choix pour l'étude de la barrière d'espèces et l'évolution de ce système d'aminacylation de manière plus générale. D'autre part, cette étude apportera de nouveaux éléments pour la compréhension des règles de l'aminacylation dans la mitochondrie. En effet, par comparaison avec les nombreuses études effectuées sur les bactériens, mais aussi d'archaea et d'eucaryotes, les connaissances sur les systèmes d'aminacylation des organelles sont encore très limitées. Cet état de fait s'explique en partie par les difficultés pour obtenir les constituants de ces systèmes d'aminacylation.

Le système d'aminacylation de la tyrosine dans la mitochondrie humaine sera étudié par des approches fonctionnelles puis structurales. L'objectif tout d'abord est de déterminer les paramètres de l'aminacylation de l'ARNt<sup>Tyr</sup> mitochondrial humain par la TyrRS homologue et de mettre à jour ses particularités. Par la suite nous chercherons à comprendre le mécanisme d'interaction entre l'ARNt<sup>Tyr</sup> et la TyrRS mitochondriale au regard des deux types de mécanismes d'interaction décrits chez la bactérie *T. thermophilus* et l'archaea *M. jannaschii*. Dans un second temps, la TyrRS sera étudiée par une approche structurale. La cristallisation et la résolution de sa structure à haute résolution seront entreprises pour comprendre le mécanisme réactionnel et mettront en exergue les divergences avec les autres systèmes de tyrosylation. Cette étude pluridisciplinaire apportera une vision globale de ce système et sera une avancée vers la compréhension des systèmes d'aminacylation mitochondriaux humains.

## Résultats et Discussion





## A. Obtention des molécules

La séquence de la TyrRS mitochondriale humaine, encore inconnue lors de l'initiation de ce projet, a été recherchée dans les banques de données génomiques humaines avant de pouvoir être clonée et exprimée. Ce travail a été présenté lors du "20<sup>th</sup> International tRNA Workshop" (Banz, Allemagne, Octobre 2003) sous la forme d'une affiche puis au "Forum des Jeunes Chercheurs" (Marrakech, Maroc, Mai 2004) lors d'une présentation orale. Un article édité dans les Actes du Congrès présente également ce travail.

Article n°1

### **Evolution des systèmes d'aminacylation spécifiques de la tyrosine**

Luc Bonnefond, Magali Frugier, Marie Sissler, Catherine Florentz, Richard Giegé et Joëlle Rudinger-Thirion

**2004**

Congrès International de Biochimie et Forum des Jeunes Chercheurs, Actes du Congrès, Enzymologie et Métabolisme, pp. 109-112

Afin de comprendre l'évolution unique et particulière des ARNt<sup>Tyr</sup>, nous avons entrepris de décrypter les règles spécifiant la tyrosylation chez l'humain. Pour réaliser ce projet, nous avons synthétisé par transcription *in vitro* les ARNt<sup>Tyr</sup> cytoplasmique et mitochondrial humains. Tandis que la TyrRS cytoplasmique humaine est connue et étudiée depuis une dizaine d'années, la TyrRS mitochondriale n'était pas annotée dans le génome humain. Nous avons identifié *in silico* la séquence codante dans les banques de données génomiques humaines. La séquence, après prédiction de son signal d'adressage vers la mitochondrie, a été clonée dans un vecteur d'expression permettant l'ajout de six résidus histidine à l'extrémité carboxy-terminale. Cette construction a permis de purifier l'enzyme par chromatographie d'affinité sur colonne Nickel. Les tests fonctionnels montrent que la TyrRS mitochondriale humaine purifiée tyrosyle le transcrit de l'ARNt<sup>Tyr</sup> homologue, mais aussi le transcrit de l'ARNt<sup>Tyr</sup> cytoplasmique.

## Evolution des systèmes d'aminocyclation spécifiques de la tyrosine

L. Bonnefond<sup>1</sup>, M. Frugier, M. Sissler, C. Florentz, R. Giegé et J. Rudinger-Thirion

<sup>1</sup>UPR 9002 du CNRS, Institut de Biologie Moléculaire et Cellulaire, 15 rue René Descartes, F-67084 Strasbourg Cedex, France, email : L.Bonnefond@ibmc.u-strasbg.fr

### Résumé

**La fidélité de la synthèse protéique est essentielle au bon déroulement de la vie cellulaire. Elle repose d'une part sur la reconnaissance spécifique des codons de l'ARN messager par les anticodons des ARN de transfert (ARNt) au niveau du ribosome, mais également sur la charge spécifique d'un ARNt par son acide aminé homologue. Cette dernière étape est catalysée par les aminoacyl-ARNt synthétases (aaRS) réparties en deux classes selon la structure de leur site catalytique. La tyrosyl-ARNt synthétase (TyrRS) est particulière parmi les aaRS : bien que présentant les motifs structuraux caractéristiques des aaRS de la classe I, elle se comporte fonctionnellement comme une aaRS de classe II. De plus, son ARNt homologue, l'ARNt<sup>Tyr</sup> présente une première paire de bases 1-72 de nature différente selon les phylae, rendant toute charge croisée impossible.**

Afin de mieux comprendre l'évolution unique et particulière des ARNt<sup>Tyr</sup>, nous avons entrepris de décrypter les règles spécifiant la tyrosylation chez l'humain. Pour réaliser ce projet, nous avons synthétisé par transcription *in vitro* les ARNt<sup>Tyr</sup> cytoplasmique et mitochondrial humains. Tandis que la TyrRS cytoplasmique humaine est connue et étudiée depuis une dizaine d'années, la TyrRS mitochondriale n'était pas annotée dans le génome humain. Nous avons identifié *in silico* la séquence codante dans les banques de données génomiques humaines. La séquence, après prédiction de son signal d'adressage vers la mitochondrie, a été clonée dans un vecteur d'expression permettant l'ajout de 6 résidus histidine à l'extrémité carboxy-terminale. Cette construction a permis de purifier l'enzyme par chromatographie d'affinité sur colonne Nickel. L'utilisation de tests fonctionnels montre que la TyrRS mitochondriale humaine purifiée tyrosyle le transcrite de l'ARNt<sup>Tyr</sup> homologue, mais aussi le transcrite de l'ARNt<sup>Tyr</sup> cytoplasmique.

**Mots clés** ARNt<sup>Tyr</sup>, évolution, identité, mitochondrie humaine, tyrosyl-ARNt synthétase

### Introduction

La synthèse protéique, étape clé de l'expression du génome, nécessite la présence de molécules adaptatrices que sont les ARN de transfert (ARNt)

qui associent à chaque codon lu sur l'ARN messager l'acide aminé correspondant. La fidélité de cette étape repose sur le déchiffrement correct du code génétique, mais aussi sur la spécificité d'aminocyclation d'un ARNt par son aminoacyl-ARNt synthétase homologue. Ce sont des éléments présents majoritairement aux extrémités de la structure en "L" de l'ARNt, appelés éléments d'identité, mais aussi au niveau de la synthétase. Ces éléments permettent l'interaction productive entre l'ARNt et la synthétase homologue (Giegé et coll., 1998). Les différents jeux d'identité sont globalement conservés au cours de l'évolution. Cette règle n'est pas retrouvée dans le système de tyrosylation. En effet, le rôle majeur joué par la première paire de bases 1-72 empêche toute charge croisée entre procaryotes et eucaryotes (Wakasugi et coll., 1998). Afin de mieux comprendre l'évolution de ce système, nous avons choisi d'étudier la réaction de tyrosylation de l'ARNt<sup>Tyr</sup> mitochondrial humain qui présente des caractéristiques à la fois procaryotique (première paire de bases G<sub>1</sub>-C<sub>72</sub>) et eucaryotique (petite région variable) (Fig. 1).

### Matériel et Méthodes

#### Transcription des ARNt<sup>Tyr</sup> *in vitro*

Les séquences correspondantes aux différents ARNt étudiés ont été clonées sous forme de "transzyme" dans le vecteur pUC119 (Fechter et coll., 1998). Les "transzymes" sont constitués successivement de la séquence promotrice optimale reconnue par la T7 ARN polymérase, de la séquence d'un ribozyme en tête de marteau et de la séquence de l'ARNt concerné. Après transcription, l'activité autocatalytique du "transzyme" libère deux molécules : le ribozyme et l'ARNt.

#### Clonage et purification de la TyrRS mitochondriale humaine

La séquence identifiée *in silico* a été amplifiée par PCR à partir d'une banque d'ADNc humain et insérée dans le vecteur d'expression pQE70 (Qiagen). Ce vecteur ajoute un marqueur poly-histidine à l'extrémité C-terminale de la protéine clonée. La protéine a été surexprimée dans une souche d'*Escherichia coli* Top10 après induction avec 1mM d'isopropyl-β-D-thiogalacto-pyranoside pendant 3h. La TyrRS mitochondriale humaine a été purifiée sur une colonne d'affinité substituée avec du Nickel.

#### Tyrosylation

La réaction de tyrosylation *in vitro* est quantifiée par l'usage de tyrosine radioactive. Des aliquotes du milieu réactionnel sont prélevées à différents temps et déposées sur des rondelles de papier Whatman 3MM. Celles-ci sont plongées dans une solution de TCA 5% froid afin de précipiter l' aminoacyl-ARNt radioactif. Après lavage, les rondelles sont séchées et déposées dans des fioles contenant du liquide de scintillation. La radioactivité est quantifiée dans un compteur à scintillation.

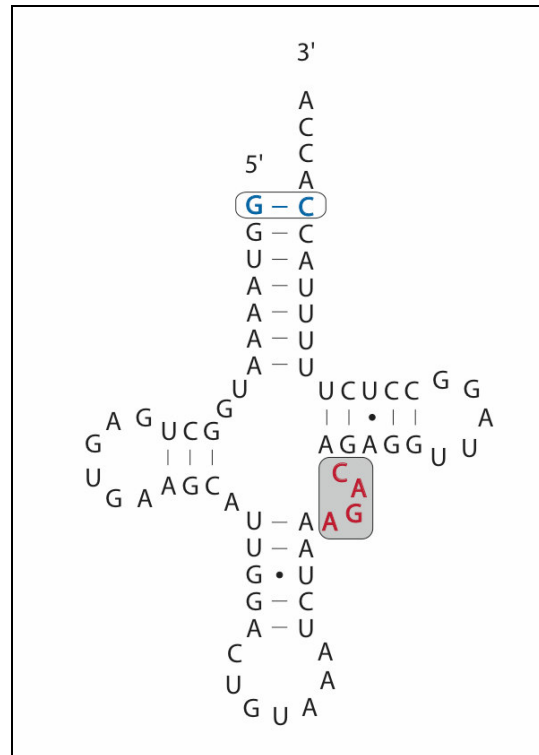
**Résultats**

**Recherche de la séquence de la TyrRS mitochondriale humaine**

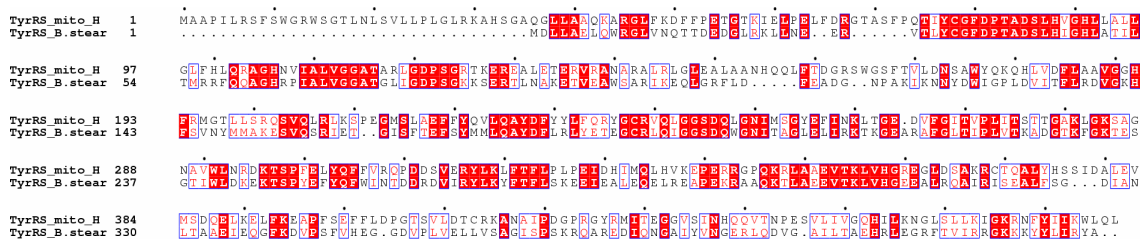
Les banques de données génomiques humaines (Ensembl humaine, euGenes, NCBI) ont été interrogées en utilisant des séquences de TyrRS procaryotiques comme sonde. Un candidat nucléique potentiel, référencé AK024057 (NEDO human cDNA sequencing project) et son équivalent protéique BAB14806, ont été identifiés. Pour valider cette séquence, celle-ci a été réinjectée comme sonde dans l'ensemble des banques de données protéiques. Les séquences extraites et le candidat potentiel ont été alignés à l'aide l'algorithme DbClustal (Fig. 2). Notre candidat s'aligne avec plusieurs séquences de TyrRS procaryotiques et présente notamment 38% avec la TyrRS de *Bacillus stearothermophilus*. La TyrRS cyto-plasmique humaine est absente de cet alignement car elle ne présente pas assez de similitudes avec l'équivalent mitochondrial.

La nature mitochondriale de cette TyrRS a été confirmée par l'identification d'un signal d'adressage à la mitochondrie à l'extrémité N-terminale. Il a été délimité grâce à l'utilisation de logiciels spécifiques (MitoProt), de logiciels de prédiction de structure secondaire (GOR IV), mais aussi à l'aide de l'alignement multiple avec d'autres TyrRS (Fig. 2). Il

englobe les 31 premiers acides aminés de la protéine. La séquence choisie pour le clonage est celle de la protéine délétée de ces 31 premiers acides aminés.



**Figure 1.** Structure secondaire du transcrite de l'ARNt<sup>Tyr</sup> mitochondrial humain. Les éléments caractéristiques des ARNt<sup>Tyr</sup> procaryotiques (entourés) et eucaryotiques (cadre gris) sont indiqués.



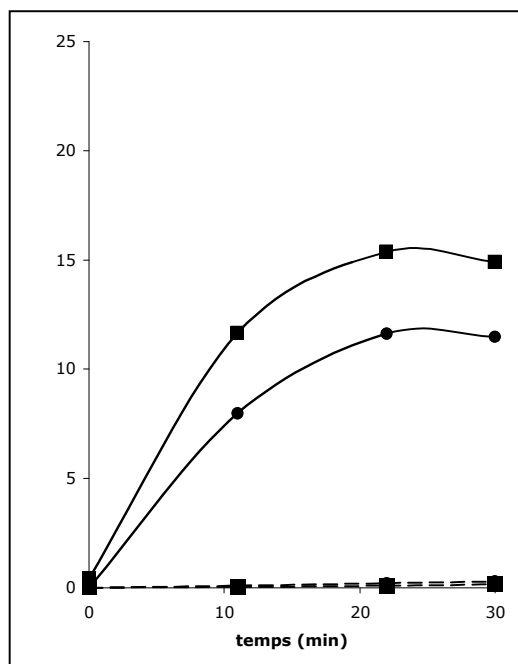
**Figure 2.** Alignement des séquences de la TyrRS mitochondriale humaine et de la TyrRS cytoplasmique de *Bacillus stearothermophilus*. Extrait de l'alignement multiple de séquences réalisé avec ClustalX.

**Paramètres cinétiques de la tyrosylation des transcrits de l'ARNt<sup>Tyr</sup> par la TyrRS mitochondriale humaine**

La spécificité d' aminoacylation de la synthétase surproduite a été vérifiée en utilisant des ARNt<sup>Tyr</sup> d'origine diverses et deux milieux réactionnels. Le milieu 1 a été préalablement mis au point pour

l' aminoacylation par la TyrRS cytoplasmique humaine (Wakasugi et coll., 1998), et le milieu 2 pour la leucylation par la LeuRS mitochondriale humaine (Sohm et coll., 2003). Dans le milieu 1, aucune charge des transcrits de l'ARNt<sup>Tyr</sup> humain par la TyrRS mitochondriale n'est détectable. L'utilisation du milieu 2 permet de tyrosyler les

transcrits de l'ARNt<sup>Tyr</sup> cytoplasmique et mitochondrial. Ceux-ci sont chargés à des taux de 10% et 20%, respectivement (Fig. 3). Cependant, ces valeurs ne tiennent cependant pas compte de la proportion de molécules d'ARNt actives présentes dans le milieu.



**Figure 3.** Plateaux de tyrosylation des deux transcrits d'ARNt<sup>Tyr</sup> cytoplasmique (●) et mitochondrial (■) humains par la TyrRS mitochondriale humaine. Les réactions réalisées dans le milieu 1 sont représentées en pointillés (---) et en trait plein (—) pour le milieu 2.

## Discussion

### Spécificité de la TyrRS

La capacité de tyrosylation du transcrit de l'ARNt<sup>Tyr</sup> mitochondrial humain par la protéine surexprimée confirme la spécificité de celle-ci. De plus, seule

l'utilisation d'un milieu spécifique aux synthétases mitochondriales permet son activité, confortant l'origine mitochondriale de la synthétase. Néanmoins, cette enzyme est également capable de tyrosiler le transcrit cytoplasmique à un taux proche de 10%. Les constantes cinétiques d'aminoacylation des transcrits cytoplasmique et mitochondrial sont proches, ce qui indique un relâchement de spécificité de la TyrRS mitochondriale humaine. Il semble que cette synthétase contraste avec les autres TyrRS qui sont, quant à elles, très spécifiques à leur phylum. Puisque la paire de bases 1-72 est de nature différente dans les deux transcrits testés, il se pourrait que l'identité tyrosine repose sur d'autres éléments qui reste à déterminer.

Les paramètres cinétiques  $K_M$  (constante de Michaelis) et  $k_{cat}$  (constante catalytique) ont été déterminés pour les deux transcrits de l'ARNt<sup>Tyr</sup> humain dans le milieu 2 ainsi que le rapport  $k_{cat}/K_M$  qui est représentatif de l'efficacité catalytique. Les valeurs sont récapitulées dans le **Tableau 1** et comparées à celles publiées pour la TyrRS cytoplasmique humaine (Jia et coll., 2003). Le transcrit de l'ARNt<sup>Tyr</sup> mitochondrial qui est le substrat homologue de la TyrRS mitochondriale présente un  $K_M$  de 3,5  $\mu\text{M}$  ce qui reflète une faible affinité de l'enzyme pour son substrat. La constante  $k_{cat}$  mesurée ( $1,7 \cdot 10^{-3} \text{ s}^{-1}$ ) est également faible. Ainsi, le rapport  $k_{cat}/K_M$  a une valeur proche de celle obtenue pour d'autres systèmes mitochondriaux, mais est environ 1000 fois inférieur à celui mesuré pour la TyrRS issue du cytoplasme humain. Étonnement, le transcrit de l'ARNt<sup>Tyr</sup> cytoplasmique présente des valeurs similaires à celle du transcrit mitochondrial.

Cependant, l'absence de modifications des transcrits peut être à l'origine d'une plus grande flexibilité de ces molécules. Celles-ci pourraient s'adapter plus facilement à la synthétase mitochondriale. (**Tableau 1**).

**Tableau 1.** Paramètres cinétiques de la tyrosylation des transcrits des ARNt<sup>Tyr</sup> cytoplasmique et mitochondrial par les TyrRS humaines.

Enzyme / Transcrit	$K_M$ ( $\mu\text{M}$ )	$k_{cat}$ ( $\text{s}^{-1}$ )	$k_{cat}/K_M$ ( $\text{s}^{-1} \cdot \mu\text{M}^{-1}$ )
<b>TyrRS cytoplasmique humaine</b>			
tRNA <sup>Tyr</sup> cytoplasmique	0,9	1,49	1,65
<b>TyrRS mitochondriale humaine</b>			
tRNA <sup>Tyr</sup> mitochondrial	3,5	$1,7 \cdot 10^{-3}$	$0,49 \cdot 10^{-3}$
tRNA <sup>Tyr</sup> cytoplasmique	2,5	$0,6 \cdot 10^{-3}$	$0,24 \cdot 10^{-3}$

### Perspectives

Afin de poursuivre l'étude fonctionnelle de la TyrRS mitochondriale humaine, il convient de déterminer l'identité de l'ARNt<sup>Tyr</sup> et de la TyrRS

mitochondriaux humains à l'aide de mutants de ces deux partenaires. Il convient aussi d'étudier la structure de l'ARNt<sup>Tyr</sup> mitochondrial qui présente

comme la plupart des ARNt mitochondriaux une structure atypique.

Parallèlement, la résolution de la structure cristallographique de la synthétase en interaction avec son ARNt homologue apportera des indications sur le mécanisme de l'interaction ARNt-synthétase.

#### **Bibliographie**

- Fechter P., Rudinger-Thirion J., Giegé R. & Théobald-Dietrich A. 1998. Ribozyme processed tRNA transcripts with unfriendly internal promoter for T7 ARN polymerase: production and activity, *FEBS Lett.* **436**, 99-103
- Giegé R., Sissler, M. & Florentz, C. 1998. Universal rules and idiosyncratic features in tRNA identity, *Nucleic Acids Res.* **26**, 5017-5035
- Jia, J., Li, B., Jin, Y. & Wang, T.P. 2003. Expression, purification, and characterisation of human tyrosyl-tRNA synthetase, *Protein Purif. & Expr.* **27**, 104-108
- Sohm B., Frugier M., Brulé H., Olszak K., Przykorska A. & Florentz C. 2003. Towards understanding human mitochondrial leucine aminoacylation identity, *J. Mol. Biol.* **328**, 995-1010
- Wakasugi K., Quinn C.L., Tao N. & Schimmel P. 1998. Genetic code in evolution: switching species-specific aminoacylation with a peptide transplant, *EMBO J.* **17**, 297-305

# HUMAN MITOCHONDRIAL TYROSYL-tRNA SYNTHETASE: EXPRESSION, PURIFICATION AND ACTIVITY

Bonnefond Luc<sup>1</sup>, Frugier Magali, Sissler Marie, Giegé Richard & Rudinger-Thirion Joëlle

<sup>1</sup>UPR 9002 du CNRS, Institut de Biologie Moléculaire et Cellulaire, 15 rue René Descartes, F-68074 Strasbourg Cedex, France, email : L.Bonnefond@ibmc.u-strasbg.fr

## INTRODUCTION

Translation fidelity is based on the mRNA codons specific recognition by the tRNA anticodons, but also on the tRNA specific aminoacylation with its cognate amino acid. This latter step is catalysed by aminoacyl-tRNA synthetases which are separated in two distinct classes according to the structure of their catalytic site [1].

The tyrosylation system is peculiar among the aminoacylation processes. The homodimeric class I tyrosyl-tRNA synthetase (TyrRS) has a mode of recognition of its cognate tRNA<sup>Tyr</sup> similar to that of class II synthetases [2]. Furthermore the tRNA<sup>Tyr</sup> overall architecture changed during evolution, going from a long variable arm in prokaryotes, to a short variable arm in eukaryotes and archaea. This evolutionary divergence is also correlated to the nature of the first base-pair in the cognate tRNA acceptor stem, G1-C72 in prokaryotes and mitochondria, and C1-G72 in eukaryotes and archaea (Fig 1).

Interestingly, the mitochondrial tRNA<sup>Tyr</sup> presents features belonging to both phyla: a typical prokaryotic G1-C72 base-pair and a short variable region, as for eukaryotic tRNA<sup>Tyr</sup> (Fig 2). To further gain insight the understanding of tRNA<sup>Tyr</sup> unique evolutionary process, we undertake the study of tyrosine identity in human mitochondria. Although human cytoplasmic TyrRS has already been largely studied (e.g. [3]), the corresponding mitochondrial enzyme has not been described so far.

Fig 1. Overall structure and identity elements of tyrosine specific tRNAs are not conserved during evolution (identity elements are highlighted in purple)

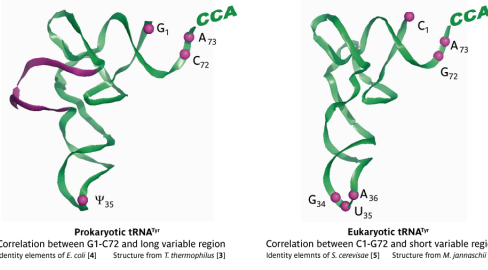
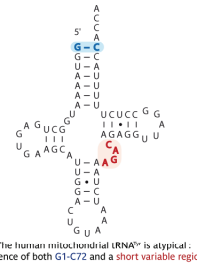


Fig 2. Secondary structure of human mitochondrial tRNA<sup>Tyr</sup> transcript



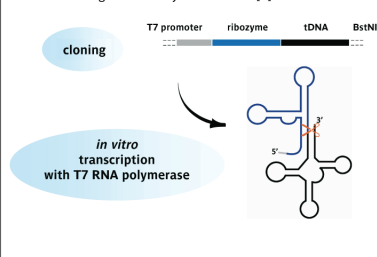
## OBJECTIVES

- Characterise the human mitochondrial TyrRS
- Determine the tyrosine identity of both human cytoplasmic and mitochondrial tRNA<sup>Tyr</sup>
- Test the tyrosylation of the human mitochondrial tRNA<sup>Tyr</sup> by the human cytoplasmic TyrRS and vice versa

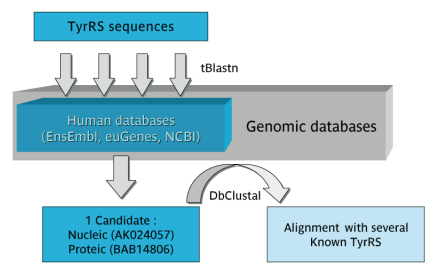
## METHODS

- Production of both cytoplasmic and mitochondrial human tRNA<sup>Tyr</sup> transcripts using the "transzyme" method [7]. The mitochondrial tRNA<sup>Tyr</sup> has an unfriendly 5' sequence for transcription by T7 RNA polymerase using the classical transcription method (Result 1).
- *In silico* search of the human mitochondrial TyrRS (Result 2).
- Analyse of the multiple alignment. Prediction of the enzyme secondary structure organisation and of the mitochondrial addressing signal (Result 3).
- Cloning, expression and purification of the human mitochondrial TyrRS (Result 4).
- Aminoacylation assays of tRNAs with the human mitochondrial TyrRS (Result 5).

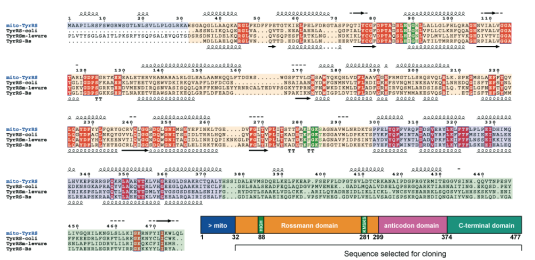
Result 1. *In vitro* synthesis of the two human tRNA<sup>Tyr</sup> transcripts using the "transzyme" method [8]



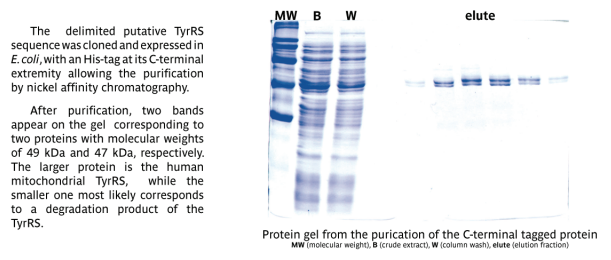
Result 2. *In silico* search of the human mitochondrial TyrRS



Result 3. Delimitation of the putative human mitochondrial TyrRS sequence



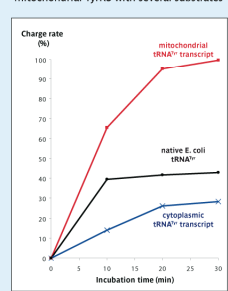
Result 4. Expression and purification of human mitochondrial TyrRS



## RESULTS

- The proteic candidate BAB14806 extracted from the human genomic databases displays both HVGH and KLGKS motifs characteristic of class I aARSs. This protein shares 38% and 37% identity with *B. stearothermophilus* and yeast mitochondrial TyrRSs, respectively. Interestingly no significant identity with the human cytoplasmic TyrRS is detected.
- The alignment of the candidate with prokaryotic TyrRSs allows the removal of the first 31 codons of the sequence selected for cloning. These codons correspond *a priori* to the mitochondrial targeting signal. Thus, the defined human mitochondrial TyrRS is composed of 447 amino-acids likely organised in three structural domains (Result 3).
- The specificity of the human mitochondrial TyrRS obtained was tested towards different substrates. Using experimental conditions recently described for human mitochondrial LeuRS [5] (50 mM HEPES-KOH pH 7.6, 25 mM KCl, 12 mM MgCl<sub>2</sub>, 2.5 mM ATP, 0.2 mg/ml BSA, 1 mM spermine), the human mitochondrial tRNA<sup>Tyr</sup> transcript is aminoacylated up to 100% by the over-expressed protein. The native *E. coli* tRNA<sup>Tyr</sup> and the human cytoplasmic tRNA<sup>Tyr</sup> transcript reach plateau levels of 40% and 30%, respectively. Determination of kinetic data are underway.

Result 5. Aminoacylation assays of the human mitochondrial TyrRS with several substrates



## CONCLUSIONS

- The over-expressed and purified protein displays tyrosylation activity. As for other aARSs (e.g. [9]), the presence of an His-tag at the C-terminal end of the protein does not hinder its correct folding.
- The fact that mitochondrial tRNA<sup>Tyr</sup> transcript is tyrosylated by the enzyme up to 100% validates the tyrosylation specificity and the mitochondrial origin of the synthetase.
- Until now, it was shown that TyrRSs are species specific. However, human mitochondrial TyrRS not only aminoacylates its cognate tRNA<sup>Tyr</sup>, but also a prokaryotic (*E. coli*) and an eukaryotic tRNA<sup>Tyr</sup> (human cytoplasmic). This suggests a looser specificity of the human mitochondrial TyrRS.

## REFERENCES

- [1] Eriani *et al.* (1990) Nature, **347**, 203-206; Cusack *et al.* (1990) Nature, **347**, 249-255
- [2] Wakasugi *et al.* (1998) EMBO J., **17**, 297-305
- [3] Yaremchuk *et al.* (2002) EMBO J., **21**, 3829-3840
- [4] Himeno *et al.* (1990) Nucleic Acids Res., **18**, 6815-6819
- [5] Fechter *et al.* (2001) Eur. J. Biochem., **268**, 761-767
- [6] Kobayashi *et al.* (2003) Nat. Struct. Biol., **10**, 425-432
- [7] Fechter *et al.* (1998) FEBS Lett., **436**, 99-103
- [8] Sohm *et al.* (2003) J. Mol. Biol., **328**, 995-1010
- [9] Xu *et al.* (2001) Protein Expr. Purif., **23**, 293-300

Figure 13. Affiche pour le "20<sup>th</sup> International tRNA Workshop" (Banz, Allemagne, 2-7 Octobre 2003)

Parallèlement à ce travail, l'AspRS mitochondriale humaine a été caractérisée au laboratoire par Aurélie Fender. La description et caractérisation conjointe de l'AspRS et de la TyrRS, ainsi que la détermination du jeu complet d'aminoacyl-ARNt synthétases mitochondriales humaines ont été publiées dans la revue "Biochemistry" (en 2005).

Article n°2

**Vers le jeu complet des aminoacyl-ARNt synthétases  
mitochondriales humaines :  
caractérisation de l'AspRS et de la TyrRS**

Luc Bonnefond\*, Aurélie Fender\*, Joëlle Rudinger-Thirion, Richard Giegé, Catherine Florentz  
et Marie Sissler ( \* co-premier auteur)

**2005**

Biochemistry, **44** : 4805-4816

La mitochondrie humaine possède une machinerie traductionnelle dédiée à la synthèse des 13 protéines mitochondriales. Tandis que les ARNt et ARNr sont produits lors de la transcription du génome mitochondrial, tous les autres facteurs nécessaires à la synthèse protéique sont synthétisés dans le cytosol puis importés. C'est notamment le cas des aminoacyl-ARNt synthétases, les enzymes qui aminoacylent les ARNt homologues avec l'acide aminé spécifique. Tandis que tous les gènes d'aaRS cytosoliques sont identifiés, seuls 9 gènes de synthétases mitochondriales ont été annotés jusqu'à présent. Dans cet article, nous présentons les gènes de l'AspRS et de la TyrRS mitochondriale humaine et la caractérisation initiale de ces enzymes. Chaque synthétase appartient à la classe de synthétases attendue, possède une organisation homodimérique et aminoacyle l'ARNt d'*Escherichia coli* correspondant ainsi que son transcrit homologue. Les gènes des autres synthétases mitochondriales ont également été identifiés, à l'exception de celui de la glutaminyl-ARNt synthétase. L'analyse de leurs séquences confirme et conforte l'idée que, à l'exception de la lysyl- et de la glycy-ARNt synthétase, les aaRS mitochondriales et cytosoliques humaines sont codées par deux jeux de gènes différents.



## Toward the Full Set of Human Mitochondrial Aminoacyl-tRNA Synthetases: Characterization of AspRS and TyrRS<sup>†</sup>

Luc Bonnefond,<sup>§</sup> Aurélie Fender,<sup>§</sup> Joëlle Rudinger-Thirion, Richard Giegé, Catherine Florentz,\* and Marie Sissler

*Department “Mécanismes et Macromolécules de la Synthèse Protéique et Cristallogénèse”, UPR 9002, Institut de Biologie Moléculaire et Cellulaire du CNRS, 15 rue René Descartes, F-67084 Strasbourg Cedex, France*

*Received November 24, 2004; Revised Manuscript Received January 19, 2005*

**ABSTRACT:** The human mitochondrion possesses a translational machinery devoted to the synthesis of 13 proteins. While the required tRNAs and rRNAs are produced by transcription of the mitochondrial genome, all other factors needed for protein synthesis are synthesized in the cytosol and imported. This is the case for aminoacyl-tRNA synthetases, the enzymes which esterify their cognate tRNA with the specific amino acid. The genes for the full set of cytosolic aaRSs are well defined, but only nine genes for mitochondrial synthetases are known. Here we describe the genes for human mitochondrial aspartyl- and tyrosyl-tRNA synthetases and the initial characterization of the enzymes. Both belong to the expected class of synthetases, have a dimeric organization, and aminoacylate *Escherichia coli* tRNAs as well as in vitro transcribed human mitochondrial tRNAs. Genes for the remaining missing synthetases were also found with the exception of glutaminyl-tRNA synthetase. Their sequence analysis confirms and further extends the view that, except for lysyl- and glycyl-tRNA synthetases, human mitochondrial and cytosolic enzymes are coded by two different sets of genes.

Aminoacyl-tRNA synthetases (aaRSs)<sup>1</sup> are key enzymes in the translation of the genetic information since they catalyze the specific attachment of each of the 20 amino acids (aa) to a cognate transfer RNA (tRNA). These enzymes have been studied equally well both at the gene and at the protein levels with most information gained on eubacterial, archæal, and eukaryal cytosolic (cyt) synthetases (reviewed in, e.g., 1–3). From a structural point of view, despite the fact that they all fulfill the same function, the 20 enzymes are partitioned into two distinct structural classes at the level of their catalytic sites with 10 enzymes in each class (4, 5). Class I enzymes bear the classical Rossmann fold (defined as an adenylic-nucleotides recognition site) that displays five parallel  $\beta$ -strands connected via  $\alpha$ -helices and the two signature motifs HIGH and KMSKS. Class II enzyme catalytic sites display an alternate folding, mainly constituted by a sheet of six antiparallel  $\beta$ -strands and three motifs of less-conserved sequences (...P... for motif 1, ...FRXE... for motif 2, and ...GXGXGER... for motif 3). From a functional point of view, rules governing specific recognition by aaRSs of their cognate amino acids (6) and tRNAs

(7–9) become well understood. From a cellular point of view, several sets of aaRSs have to coexist in the same eukaryotic cell to achieve protein synthesis within both cytosol and organelles. These three aspects raise questions on the distinguishing characteristics of mitochondrial enzymes and on the evolutionary history of their genes.

For mammals, and more specifically for humans, the genes for the 20 cytosolic enzymes are known, but limited information on mitochondrial aaRSs (mt-aaRSs) is available (10). The human mitochondrial genome itself codes for only 13 proteins (all subunits of the inner mitochondrial membrane respiratory chain complexes), 2 rRNAs, and 22 tRNAs (11), so that mt-aaRSs are necessarily nuclearly encoded and imported.

Most of the human mitochondrial aminoacylation activities were detected long ago, but genomic information on enzymes became available only recently. To our knowledge, sequence information and basic structural data on genes are available for mt-GlyRS (12, 13) and mt-HisRS (14, 15). Further, the genes of mt-IleRS (16, 17), mt-PheRS (18), mt-LysRS (19), mt-LeuRS (20), mt-TrpRS (21), mt-SerRS (22), and mt-MetRS (23) have been identified and cloned, and the corresponding proteins have been characterized. Most of these mitochondrial enzymes (HisRS, IleRS, LeuRS, MetRS, PheRS, SerRS, and TrpRSs) are coded by different genes than the corresponding cytosolic synthetases. Only 2 out of the 9 previously characterized mitochondrial enzymes are encoded by the same gene than the cytosolic enzyme. Cyt- and mt-GlyRSs are generated from two translation initiation sites on the same gene, leading to one enzyme with a mt-targeting signal and a second with cytosolic location (12,

<sup>†</sup> This work was supported by CNRS (Centre National de la Recherche Scientifique), ULP (Université Louis Pasteur, Strasbourg), and AFM (Association Française contre les Myopathies).

\* To whom correspondence should be addressed: phone, 33 3 88 41 70 59; fax, 33 3 88 60 22 18; e-mail, C.Florentz@ibmc.u-strasbg.fr.

<sup>§</sup> These authors contributed equally to the work.

<sup>1</sup> Abbreviations: tRNA, transfer RNA; aaRS, aminoacyl-tRNA synthetase with aa for amino acid (for individual aaRS, aa is given in the three letter code, e.g., AspRS for aspartyl-tRNA synthetase); mt, mitochondrial; cyt, cytosolic; EMAP, endothelial-monocyte-activating polypeptide.

13). In the case of LysRS, an alternative mRNA splicing pathway allows the insertion, or not, of the mt-targeting signal leading to two mature enzymes differing only by a few residues at their N-terminus (19).

To expand the knowledge on human mt-aARSs, we report here the identification and cloning of the genes coding for mt-AspRS and mt-TyrRS as well as the initial characterization of the two enzymes. Aminoacylation systems specific for aspartate or tyrosine from numerous organisms have already been investigated biochemically and structurally (reviewed in 24, 25). AspRSs are typical class II aARSs (they belong to subclass IIb), with a conserved modular architecture composed of a C-terminal active site domain linked to an N-terminal anticodon-binding domain by a short hinge. This organization is retained in the three domains of life. However, eukaryal AspRSs are distinguished from the others by an N-terminal extension involved in tRNA binding (26) and/or multisynthetase complex formation (27, 28). Eubacterial AspRSs have an insertion module within the active site domain and a C-terminal extension. TyrRSs belong to class I aARSs but in an ambiguous way. They share the typical class I sequence motifs, but their oligomeric structure is dimeric (a typical class II aARS characteristic), and their mode of cognate tRNA recognition is similar to that of class II aARSs (e.g., 29, 30). Each monomer of TyrRS contains an N-terminal Rossmann-fold catalytic domain followed by an anticodon-recognition domain, which is smaller in archæobacteria as compared to the other phylæ (31). In humans, cyt-TyrRS can be split into two fragments with distinct cytokine activities (32).

In an attempt to uncover the full set of genes for human mt-aARSs, we have further taken advantage of the recent throughputs in human genome sequencing, have extensively screened protein, EST (expressed sequence tag), cDNAs, and genomic databases, and have identified potential open reading frames for 8 of the 9 missing genes.

## MATERIALS AND METHODS

**Materials.** Native tRNA<sup>Asp</sup> from *Escherichia coli* was a kind gift from D. Kern (Strasbourg), and native tRNA<sup>Tyr</sup> from *E. coli* was from Subriden (Rolling Bay, WA). Purified oligonucleotides were from Proligo (Boulder, CO). L-[<sup>3</sup>H]-aspartic acid (37 Ci/mmol) and L-[<sup>3</sup>H]-tyrosine (49 Ci/mmol) were from Amersham (Sweden). Restriction enzymes (*Bam*HI, *Bgl*III, *Bst*NI, *Hind*III, and *Pst*I) and T4 polynucleotide kinase were from New England Biolabs (U.K.). T4 DNA ligase was from Qbiogen (France). T7 RNA polymerase was purified from an overproducing strain as described (33). Dynazyme EXT polymerase was from Finnzymes (Finland).

**Cloning Procedures.** The same method was used to clone the genes of human mt-AspRS and mt-TyrRS. Genes were amplified from a human cDNA bank (gift from F. Martin, Strasbourg) using Dynazyme EXT polymerase. Primers for PCR were designed with restriction sites at their 5'-ends. Fresh PCR products were directly used for cloning into pCR2.1 following the TA Cloning Kit protocol (Invitrogen, Carlsbad, CA). After transformation of One Shot TOP10 *E. coli* cells, plasmidic DNA was purified, sequenced, and digested with the corresponding restriction enzymes. The fragment containing the expected gene was purified by gel electrophoresis, eluted, and inserted into the expression vector

pQE70 (Qiagen, Germany) leading to pQE70-mt-AspRS and pQE70-mt-TyrRS plasmids.

For pQE70-mt-AspRS, primers designed for the PCR were as follows: 5'-AGATCTGTTGTCCGGACCAACACATGTGGA-3' (5'-primer) and 5'-AGATCTATGAGCTCTTTCTGCTTTGGAGTC-3' (3'-primer), both including a *Bgl*III restriction site. The PCR was performed for 2 min at 95 °C followed by 30 cycles of 30 s at 95 °C, 45 s at 52 °C, and 2 min at 72 °C and finalized with 10 min at 72 °C. Fifty nanograms of cDNA matrix and 2 U of polymerase were used along with 1 μM of each primer. Correct orientation of the gene within the pQE70 expression vector was verified by restriction fragment mapping.

For pQE70-mt-TyrRS, primers designed for the PCR were 5'-GGATCCACACTCGGGCGCTCAGGGGTTACTGG-3' (5'-primer including a *Bam*HI restriction site) and 5'-AGATCTCAACTGAAGCCATTTTATAATG-3' (3'-primer including a *Bgl*III restriction site). PCR conditions were 5 min at 95 °C followed by 25 cycles of 1 min at 95 °C, 1 min at 55 °C, and 1 min at 70 °C and finalized with 10 min at 70 °C. Two micrograms of cDNA matrix and 2 U of polymerase were used along with 1 μM of each primer. Plasmid pCR2.1-mt-TyrRS DNA was digested with *Bam*HI and *Bgl*III restriction enzymes to release the *mt-TyrRS* gene containing fragment and digested further with *Pst*I to eliminate other fragments of the same size.

**Expression of the Cloned Genes.** TOP10 strains transformed by plasmids pQE70-mt-AspRS or pQE70-mt-TyrRS were grown in a 200 mL culture of LB medium containing 50 μg/mL ampicillin and 2% glucose overnight at 37 °C. After dilution to 1:10 with 1800 mL of fresh medium, cells were grown at 37 °C for 2 h until an OD<sub>600</sub> of 0.7 was reached. The medium was then replaced by LB medium containing 50 μg/mL ampicillin and 500 or 10 μM IPTG, respectively, and the incubation continued overnight at 25 °C. The cells were harvested by centrifugation at 4000g for 10 min (4 °C).

**Purification.** TOP10-pQE70-mt-AspRS and TOP10-pQE70-mt-TyrRS cell pellets were resuspended in 45 mL of buffer A (50 mM NaH<sub>2</sub>PO<sub>4</sub> pH 7.5, 300 mM NaCl, 20 mM imidazole, 10 mM β-mercaptoethanol, and 10% glycerol) before sonication (Ultrasons Annemasse, France) on ice at 120 V for 40 s, 7 times. Lysed cells were then subjected to centrifugation at 30 000g for 45 min at 4 °C. If necessary, the supernatant was filtered through a 0.45 μm MillexHA filter (Millipore, France) before loading onto a 2 mL Ni-NTA resin column (Qiagen) equilibrated in buffer A. After the column was washed with 50 mL of buffer A, mt-AspRS and mt-TyrRS were eluted with a 30 mL linear gradient of buffer A to B (buffer A adjusted to 500 mM imidazole). The enzyme-containing fractions, detected by the OD<sub>280</sub> profile and a SDS-PAGE analysis, were pooled and dialyzed overnight at 4 °C against 1 L of buffer C (50 mM KH<sub>2</sub>PO<sub>4</sub>/K<sub>2</sub>HPO<sub>4</sub> pH 7.5, 150 mM KCl, 1 mM EDTA, 10 mM β-mercaptoethanol, and 50% glycerol) and stored at -20 °C. Protein concentrations were determined either from OD<sub>280</sub> using the theoretical extinction coefficient calculated with ProtParam from ExPASy tools ( $\epsilon = 43\,540\text{ M}^{-1}\cdot\text{cm}^{-1}$  for mt-AspRS and  $\epsilon = 39\,520\text{ M}^{-1}\cdot\text{cm}^{-1}$  for mt-TyrRS) or from the Bradford method (Bio-Rad) using bovine serum albumin as reference.

**Analytical Procedures.** Gel filtration chromatography was performed at 4 °C using a 15 mL Bio-Prep SE-100/17 column (Bio-Rad) equilibrated with buffer 1 (50 mM NaH<sub>2</sub>PO<sub>4</sub> pH 7.5 and 300 mM NaCl) for mt-AspRS or buffer 2 (50 mM HEPES–NaOH pH 7.5, 10 mM MgCl<sub>2</sub>, 150 mM KCl, 1 mM DTE, and 20% glycerol) for mt-TyrRS. The column was calibrated with a set of standard proteins (molecular weights of 1.35, 17, 44, 158, and 670 kDa, Bio-Rad). Enzyme samples of 200 μL [at 0.6 mg/mL for mt-AspRS and 2 mg/mL (HEPES–NaOH buffer) for mt-TyrRS] were applied to the column. Molecular weights (MW) were calculated from log(MW) versus elution volume plots.

**Cloning and in Vitro Transcription of tRNAs.** Clones containing genes for human mt-tRNA<sup>Asp</sup> and mt-tRNA<sup>Tyr</sup> were obtained by hybridization of nine and eight overlapping phosphorylated oligonucleotides, respectively, ligation into *Hind*III and *Bam*HI sites of plasmid pTFMA and pUC119, respectively, and transformation into *E. coli* TG1 cells. Both synthetic genes contain a T7 RNA polymerase promoter followed by a hammerhead ribozyme sequence and the tRNA sequence (34). A *Bst*NI site, coincidental with the 3'-end of the tRNA sequence, allowed the synthesis of tRNAs ending with the expected CCA-3'-sequence. Transcription was performed according to established procedures (34, 35). The hammerhead ribozyme self-cleaves the phosphodiester linkage directly upstream nucleotide 1 and liberates the tRNA. After phenol extraction, the transcripts were purified on denaturing PAGE, electro-eluted, and ethanol-precipitated. The concentration of transcripts was determined spectrophotometrically with 40 (μg/mL)/cm for 1 OD<sub>260</sub> unit.

**Aminoacylation Assays.** Aminoacylation assays were performed as described (36). Aspartylation and tyrosylation conditions were 50 mM HEPES–KOH pH 7.6, 25 mM KCl, 12 mM MgCl<sub>2</sub>, 2.5 mM ATP, 0.2 mg/mL BSA, 1 mM spermine, 32 μM [<sup>3</sup>H]aspartate (208 GBq/mmol) or 10 μM [<sup>3</sup>H]tyrosine (3 TBq/mmol) respectively, and adequate amounts of transcripts (0.04–1.6 μM) and enzymes (10–150 nM). Transcripts were renatured by heating at 60 °C for 90 s in water and slow cooling to room temperature before the aminoacylation was performed at 25 °C (aspartylation) or 37 °C (tyrosylation). The lower temperature chosen for aspartylation was required to obtain reproducible results in aminocylation. It corresponds to a compromise between optimal enzymatic activity and tRNA stability as was previously shown for in vitro transcribed mt-tRNA<sup>Lys</sup> (35). This is likely a more general situation since human mt-tRNAs are encoded by either of the two DNA strands of strongly biased nucleotide composition and thus are either “heavy” and stable (tRNA<sup>Tyr</sup>) or “light” and thermodynamically weak (tRNA<sup>Asp</sup>) (37). After different incubation times, aliquots were removed, spotted on Whatman 3MM paper, and precipitated in 5% trichloroacetic acid. Incorporation of radioactive amino acid was measured by liquid scintillation counting. Plateau levels of charging were 60% for mt-tRNA<sup>Asp</sup> and 90% for mt-tRNA<sup>Tyr</sup>. Kinetic parameters  $k_{cat}$  and  $K_M$  were derived from Lineweaver–Burk plots. Displayed data represent an average of three independent experiments.

**In Silico Methods.** Genes for human aaRSs have been searched in selected databases (Ensembl Human database, NCBI protein, EST, cDNAs, and genomic databases, and euGenes), using entire or partial sequences of known aaRSs

(of a given specificity) from other organisms as queries. Programs were (i) standard nucleotide–nucleotide Blast (Blastn) which compares nucleotide sequence queries against the NCBI nucleotide databases, (ii) tBlastn, which compares a protein sequence against NCBI nucleotide database translated into all 6 reading frames, and (iii) the standard protein–protein Blasts, Blastp and Ballast. The first two methods lead to nucleic acid sequences or accession numbers (DNA, cDNA, or mRNA), and the third method leads to protein sequences or accession numbers.

**Programs and Web sites are** 3DCoffee (<http://igs-server.cnrs-mrs.fr/Tcoffee/>), Ballast (<http://igbmc.u-strasbg.fr:8080/ballast.html>), Blast (<http://www.ncbi.nlm.nih.gov/blast/>), BoxShade ([http://www.ch.embnet.org/software/BOX\\_form.html](http://www.ch.embnet.org/software/BOX_form.html)), DbClustal (<http://igbmc.u-strasbg.fr:8080/DbClustal/dbclustal.html>), Ensembl Human database ([http://www.ensembl.org/Homo\\_sapiens/](http://www.ensembl.org/Homo_sapiens/)), ESPript (<http://prodes.toulouse.inra.fr/ESPrpt/ESPrpt/>), euGenes (<http://iubio.bio.indiana.edu:8089/>), ExpASY (Expert Protein Analysis System) Proteomics tools (<http://www.expasy.org/tools/>), MitoProt (<http://ihg.gsf.de/ihg/mitoprot.html>), NCBI databases (<http://www.ncbi.nlm.nih.gov/>), Prediction servers at CBS (<http://www.cbs.dtu.dk/services/>), PredictProtein (<http://www.embl-heidelberg.de/predictprotein/predictprotein.html>), Predotar (<http://genoplante-info.infobiogen.fr/predotar/predotar.html>), and TargetP (<http://www.cbs.dtu.dk/services/TargetP/>).

## RESULTS

**In Silico Characterization of Mitochondrial Aspartyl- and Tyrosyl-tRNA Synthetases.** Screening of genomic databases with a set of known synthetases of same specificity pinpointed reliable candidates for both human mt-AspRS and mt-TyrRS genes. Their respective gene names are *HSM804946* (nucleic acid accession number AL833633) and *FLJ13995* (nucleic acid accession number AK024057 and protein accession number BAB14806). The corresponding translated sequences, named herein mt-AspRS and mt-TyrRS, are displayed in Figures 1 and 2 (first lines). Arguments in support of the two sequences to be the expected enzymes are the following. Both candidate sequences hit with high scores a cluster of aaRSs of same specificity (a selection of these is presented in Figures 1 and 2). Further, catalytic site motifs of aaRSs (4, 5) are present. For mt-TyrRS, HVGH and KLGKS consensus sequences mimic the class I specific HIGH and KMSKS sequences; for mt-AspRS, motif 1 (...P...), motif 2 (...YRDE...), and motif 3 (...GGIAL-GLDRILICLV....) defining class II aaRSs are present. Additionally, both sequences contain an N-terminal mitochondrial-targeting sequence as supported by two predictive programs (Predotar and MitoProt). These programs calculate the N-terminal protein region that can support a mt-targeting sequence and predict possible cleavage sites for the tag after import into the mitochondria (see “predicted tag cleavage positions” in Figures 1 and 2). Probabilities to be mitochondria addressed were found as 0.816 and 0.998 for mt-AspRS and 0.944 and 0.974 for mt-TyrRS. These scores are indicative of mitochondrial sequences and allowed us to get rid of false positive candidate sequences which all had scores below 0.05.

The gene for human mt-AspRS codes for a total of 645 aa from which 47 are predicted to correspond to the mt-



# AspRS

## Tag cleavage position

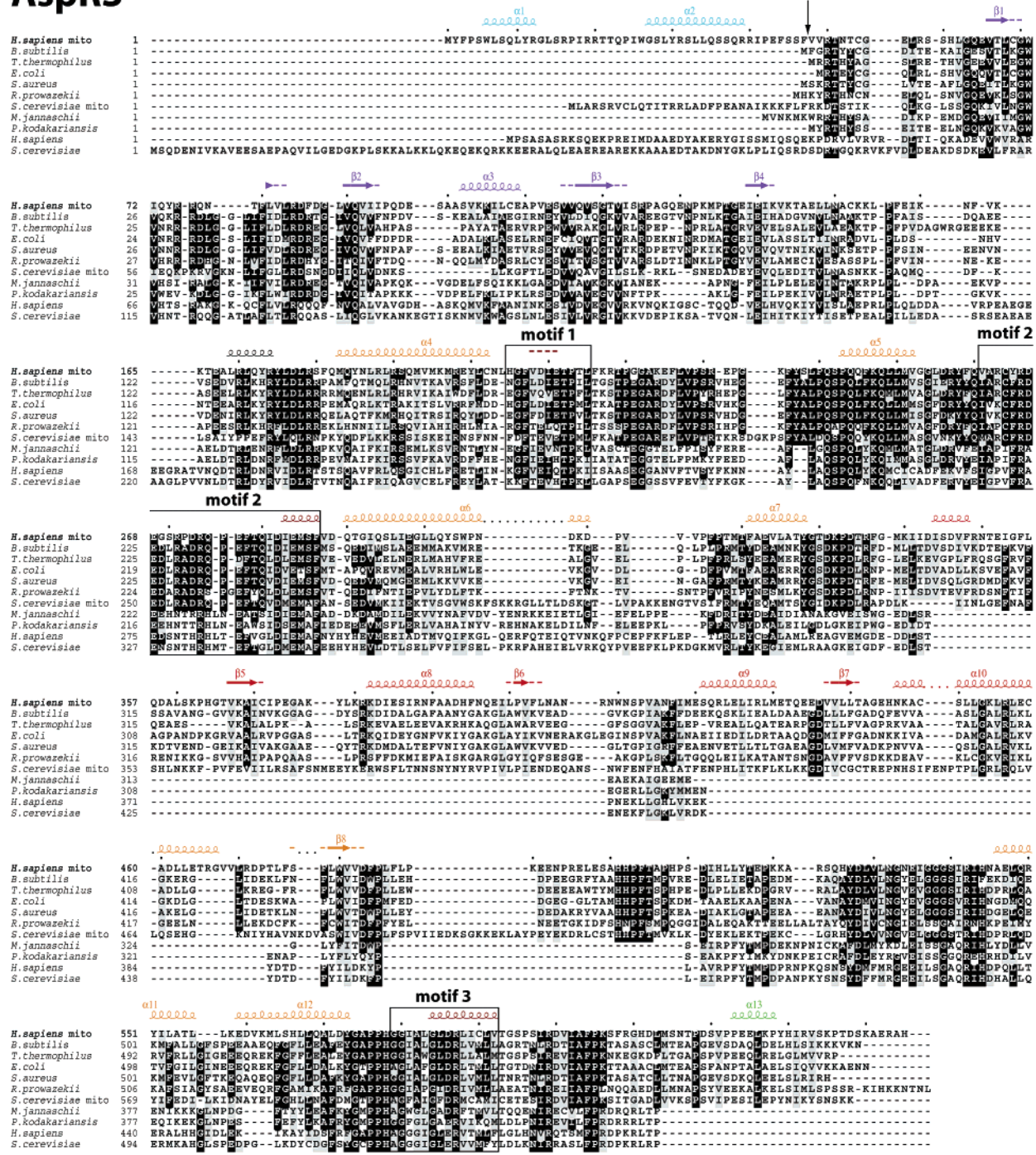


FIGURE 1: Sequence organization of human mitochondrial AspRS. Multiple sequences alignment of human mt-AspRS (in bold) with a selection of known AspRSs has been established using 3DCoffee program (70) and shaded with BoxShade. Notice that sequences have been reordered in accordance with their hit score toward human mt-AspRS. Black and gray boxes highlight, respectively, strictly identical and similar residues. Human mt-AspRS structure elements displayed above the sequence ( $\alpha$ -helices and  $\beta$ -strands) have been predicted using PredictProtein (71) and drawn with ESPript (structure 2D) (72). For the color code, see Figure 5. The three signature motifs specifying class II aaRSs are shown. Accession numbers of sequences are *Bacillus subtilis*, P36419; *Escherichia coli*, P21889; *Staphylococcus aureus*, Q99TL9; *Rickettsia prowazekii*, Q9ZE17; *Saccharomyces cerevisiae* mt, NP\_015221; *Thermococcus kodakaraensis*, Q52428; *Methanococcus jannaschii*, B64494; *Homo sapiens*, AAH00629; *Saccharomyces cerevisiae*, P04802. Cleavage position for the mitochondrial tag is indicated by an arrowhead.

targeting sequence, thus leading to a mature enzyme of 598 aa. Multiple alignment with related enzymes (Figure 1) reveals 36–43% identity of the full-length sequence with bacterial sequences, 32% with the mitochondrial sequence from the lower eukaryote *S. cerevisiae*, and below 23% with enzymes from archaea and cytosol of eukaryotes, including human cyt-AspRS. Alignment also shows that human mt-

AspRS possesses strictly conserved residues found in all known AspRS sequences (24). These include residues involved in ATP binding and tRNA binding. Residues involved in amino acid binding include those typical for class II aaRSs and those specific for aspartic acid recognition. Further, the alignment illustrates the already known strong conservation of AspRSs from different phylae. Clusters of



# TyrRS

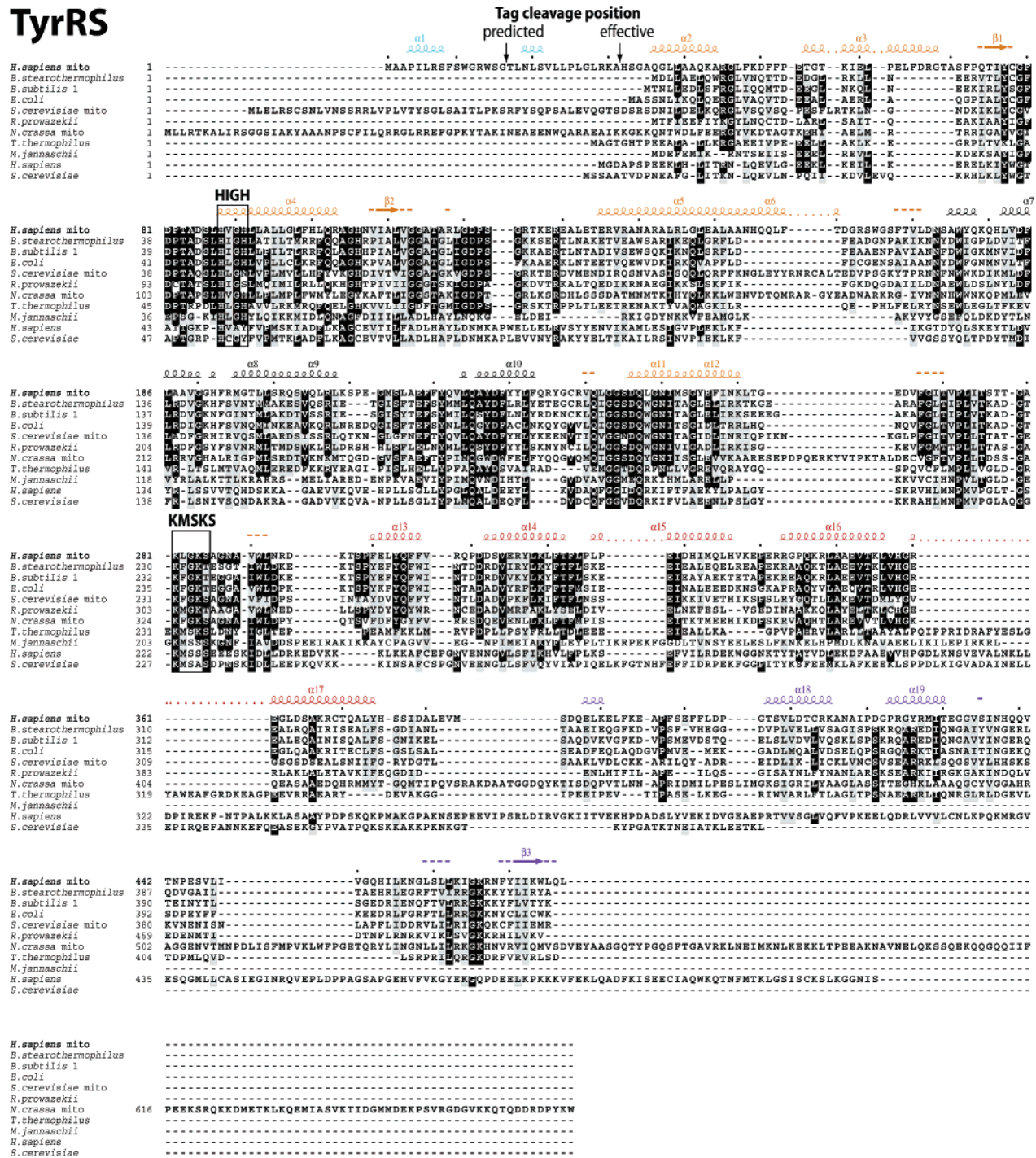


FIGURE 2: Sequence organization of human mitochondrial TyrRS. Multiple sequences alignment of human mt-TyrRS (in bold) with a selection of known TyrRSs has been established as for mt-AspRS (see legend of Figure 1 for details). The two signature motifs HIGH and KMSKS specifying class I aaRSs are framed. Accession numbers of sequences are *Bacillus stearothermophilus*, P00952; *Bacillus subtilis*, P22326; *Escherichia coli*, P00951; *Saccharomyces cerevisiae* mt, P48527; *Rickettsia prowazekii*, Q9ZCZA; *Neurospora crassa* mt, P12063; *Thermus thermophilus*, P83453; *Methanococcus jannaschii*, Q57834; *Homo sapiens* cyt, P54577; *Saccharomyces cerevisiae* cyt, P36421. Predicted and effective tag cleavage positions (see text) are indicated by arrowheads.

strongly conserved amino acids are visible (black squares) all along the sequence except within domains corresponding to the eubacterial insertion (from residue 328 to 456) and C-terminal extension (from residue 607 to 645).

The gene for human mt-TyrRS encodes a protein of 477 aa including a predicted 16 amino acid long mt-tag (Figure 2, upper line). Alignments with known TyrRS sequences from all domains of life (of which some examples are shown in Figure 2) reveal 25–40% identity with a panel of

eubacterial TyrRSs, 33 and 34% with yeast and *Neurospora crassa* mt-TyrRS sequences, and less than 25% with the enzymes from archæa and eukaryotes. Noticeable is the case of the human cyt-TyrRS that presents no significant identity with its homologous mitochondrial enzyme. From this alignment, it further appears that the greatest homologies between all TyrRSs are exclusively contained within the N-termini, whereas C-terminal domains vary in size as well as in sequence. The very low sequence conservation is further

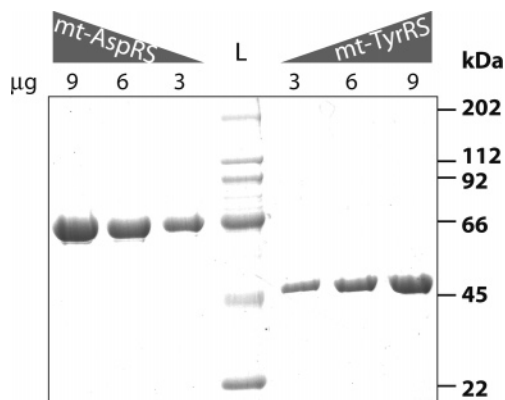


FIGURE 3: SDS-PAGE analysis of overexpressed and purified mt-AspRS and mt-TyrRS. Samples of 3, 6, and 9  $\mu\text{g}$  of enzyme preparations were analyzed on a 10% SDS-PAGE, and the protein bands were visualized by staining with Coomassie blue. Mt-AspRS is on the left side of the gel, and mt-TyrRS is on the right side of the gel. L stands for molecular weight ladder (standard mixture for molecular weights 22–202 kDa from Sigma).

illustrated by the difficulty to find residues conserved in all phylae. This is opposite to the case of AspRSs described above.

**Biochemical Characterization of Human Mitochondrial Aspartyl- and Tyrosyl-tRNA Synthetases.** The mature versions of the two enzymes were cloned into pQE70 vectors and expressed in *E. coli* strain TOP 10. As mentioned above, a theoretical cleavage site of the mt-targeting sequence can be predicted in silico for each enzyme. However, in the case of mt-TyrRS, we have designed a new cleavage site (see “effective tag cleavage position” in Figure 2). This was based on the fact that multiple alignments with bacterial TyrRS sequences suggest a larger targeting signal than that predicted in silico. Accordingly, a recombinant enzyme starting at position 32 (instead of 17), upstream a structurally predicted helix, was designed. As a result of the cloning procedure, additional residues were introduced at both ends of the molecule: four at the N-terminus and eight at the C-terminus (including six histidines). As a consequence, the recombinant human mt-TyrRS contains 458 aa, and its theoretical molecular weight is 51 302 Da. In the case of mt-AspRS, the mt-targeting sequence is of 47 aa, leading to a mature enzyme of 598 aa. For cloning purposes, six and eight aa (including six histidines) have been inserted at the N- and C-terminal parts, respectively. The recombinant human mt-AspRS contains 612 aa, with a theoretical molecular weight of 69 627 Da.

The two proteins were overexpressed in *E. coli* under similar conditions and purified by affinity chromatography on a nickel column. Human mt-TyrRS was eluted at 100 mM and mt-AspRS at 170 mM imidazole. The two proteins are pure to more than 95% as estimated on a Coomassie blue stained SDS-polyacrylamide gel (Figure 3). To be noticed is the low solubility of both enzymes in a variety of buffers (Tris-HCl, MES-NaOH, HEPES-NaOH, sodium cacodylate, potassium phosphate), and this at different pH and in a large range of ionic strength values. In short, mt-AspRS is only significantly soluble in phosphate buffers ( $\leq 1$  mg/mL), while mt-TyrRS is more soluble both in phosphate and in HEPES-NaOH buffers (to a maximal concentration of 2 and 4 mg/mL, respectively). Detailed data on solubility properties will be published elsewhere.

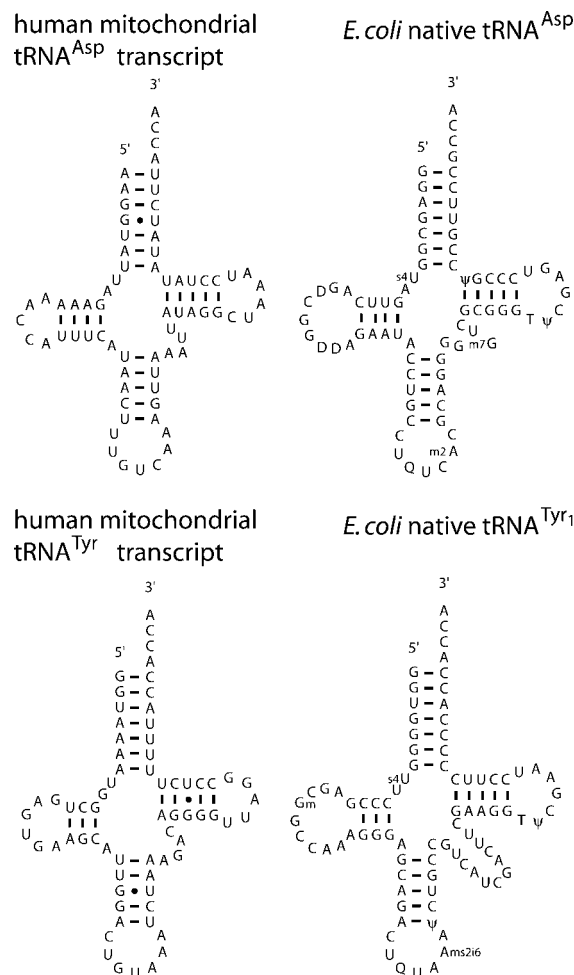


FIGURE 4: Secondary structures of tRNA<sup>Asp</sup> and tRNA<sup>Tyr</sup>, substrates of human mt-AspRS and mt-TyrRS, respectively. tRNAs are either in vitro transcripts based on human mt sequences (73) (left part) or native tRNAs from *E. coli* (right part). Names of modified bases are according to ref 74.

Table 1: Kinetic Parameters for tRNA Aminoacylation by Human mt-AspRS and mt-TyrRS

tRNA	$K_M$ ( $\mu\text{M}$ )	$k_{cat}$ ( $10^{-3} \text{ s}^{-1}$ )	$k_{cat}/K_M$ ( $10^{-3} \text{ s}^{-1} \cdot \mu\text{M}^{-1}$ )
mt-AspRS Enzyme			
<i>E. coli</i> native tRNA <sup>Asp</sup>	0.15 $\pm$ 0.05	133 $\pm$ 23	887
human mt-tRNA <sup>Asp</sup> transcript	0.13 $\pm$ 0.07	10 $\pm$ 3	77
mt-TyrRS Enzyme			
<i>E. coli</i> native tRNA <sup>Tyr</sup>	0.52 $\pm$ 0.2	210 $\pm$ 70	400
human mt-tRNA <sup>Tyr</sup> transcript	4.8 $\pm$ 1.2	46 $\pm$ 3	10

To evaluate the oligomeric status of both enzymes, gel filtration chromatographies were performed. Mt-AspRS and mt-TyrRS migrate with apparent molecular weights of 143 and 130 kDa, respectively, suggesting that they are dimers (results not shown). This is in agreement with their belonging to subclasses IIb and Ib, respectively, and does not depart from what was established for aaRSs of same specificities but from different organisms.

The capacity of both synthetases to aminoacylate tRNAs has been verified on purified *E. coli* tRNA<sup>Asp</sup> and tRNA<sup>Tyr</sup> as well as on in vitro transcripts of human mitochondrial tRNA<sup>Asp</sup> and tRNA<sup>Tyr</sup> (Figure 4, Table 1). *E. coli* tRNAs are classical substrates to test for mitochondrial activities and were shown to replace well the hardly accessible native



Table 2: Human Cytosolic and Mitochondrial Aminoacyl-tRNA Synthetases in Databases<sup>a</sup>

class I aaRSs				class II aaRSs			
aaRSs	gene name	nucl acc no.	prot acc no.	aaRSs	gene name	nucl acc no.	prot acc no.
<i>ArgRS</i>	cyt <i>RARS</i>	NM_002887	NP_002878	<i>AlaRS</i>	cyt <i>AARS</i>	NM_001605	NP_001596
	mt <i>LOC57038</i>	NM_020320	NP_064716		mt <i>LOC221410</i>	XM_291190	XP_291190
<i>CysRS</i>	cyt <i>CARS</i>	NM_001751	NP_001742	<i>AsnRS</i>	cyt <i>NARS</i>	NM_004539	NP_004530
	mt <i>FLJ12118</i>	NM_024537	NP_078813		mt <i>FLJ32418</i>	AK056980	
<i>GlnRS</i>	cyt <i>QARS</i>	NM_005051	NP_005042	<i>AspRS</i>	cyt <i>DARS</i>	NM_001349	NP_001340
	mt not found	not found	not found		mt <i>HSM804946</i>	AL833633	
<i>GluRS</i>	cyt <i>EPRS</i>	NM_004446	NP_004437	<i>GlyRS</i>	cyt <i>GARS</i>	NM_002047	NP_002038
	mt <i>KIAA1970</i>	BC040013	AAH40013		mt <i>GARS</i>	NM_002047	
<i>IleRS</i>	cyt <i>IARS</i>	NM_013417	NP_038203	<i>HisRS</i>	cyt <i>HARS</i>	NM_002109	NP_002100
	mt	D28500	BAA95147		mt <i>HARSL</i>	NM_012208	NP_036340
<i>LeuRS</i>	cyt <i>LARS</i>	D84223	BAA95667	<i>LysRS</i>	cyt <i>KARS</i>	NM_005548	NP_005539
	mt <i>LARS2</i>	NM_015340	NP_056155		mt	AF285758	AAG30114
<i>MetRS</i>	cyt <i>MARS</i>	NM_004990	NP_004981	<i>PheRS</i>	cyt $\alpha$ <i>FARSL</i>	NM_004461	NP_004452
	mt <i>BC009115</i>	NM_138395	NP_612404		cyt $\beta$ <i>PheHB</i>	NM_005687	NP_005678
					mt <i>FARSI</i>	NM_006567	NP_006558
<i>TrpRS</i>	cyt <i>WARS</i>	NM_004184	NP_004175	<i>ProRS</i>	cyt <i>EPRS</i>	NM_004446	NP_004437
	mt <i>WARS2</i>	NM_015836	NP_056651		mt	NM_152268	NP_689481
<i>TyrRS</i>	cyt <i>YARS</i>	NM_003680	NP_003671	<i>SerRS</i>	cyt <i>SARS</i>	NM_006513	NP_006504
	mt <i>FLJ13995</i>	AK024057	BAB14806		mt <i>SARSM</i>	NM_017827	NP_060297
<i>ValRS</i>	cyt <i>VAR2</i>	NM_006295	NP_006286	<i>ThrRS</i>	cyt <i>TARS</i>	NM_003191	NP_689508
	mt <i>FLJ20504</i>	AK000511			mt	BC000541	AAH00541

<sup>a</sup> Cytosolic (cyt) and mitochondrial (mt) aaRSs are sorted alphabetically according to their belonging to class I (left part) or class II (right part) (4, 5). Newly “in silico” identified aaRSs are in boldfaced italic font. Only the gene for GlnRS was not found. Accession numbers (acc. no.) are compatible with NCBI databases.

human mt-tRNAs (e.g., 18–20, 23). Aminoacylation properties of both enzymes were tested under the same conditions except that aspartylation was measured at 25 °C and tyrosylation at 37 °C. The nonphysiological lower temperature used for aspartylation was required to stabilize the structure of the in vitro transcribed substrate. The two recombinant enzymes are catalytically active, and both recognize and charge specific *E. coli* tRNAs with similar efficiencies (estimated by the ratio  $k_{cat}/K_M$ ). Both enzymes aminoacylate also their cognate substrates produced by in vitro transcription, albeit with a 10- to 40-fold reduced efficiency than they charge *E. coli* native tRNAs.

*Databank Mining for Completion of the Human Mitochondrial Aminoacyl-tRNA Synthetases Set.* Cyberscreening within databases for the nine missing human mt-aaRS genes was done similarly to the search of human mt-AspRS and mt-TyrRS. Already known aaRS sequences of a given specificity were used as queries, and a search for a mt-targeting sequence was performed. Candidate sequences were found for eight mt-aaRSs, that is, AlaRS, ArgRS, AsnRS, CysRS, GluRS, ProRS, ThrRS, and ValRS. As anticipated, no gene for mt-GlnRS was found. The newly identified enzymes contain either of the consensus sequences defining the catalytic site of a synthetase. However, two particular cases have to be emphasized. Mt-ArgRS has an “-MKTR” motif instead of “KMSKS”, a situation already reported for other ArgRSs (38), and human mt-AlaRS is deprived of motifs 1 and 2, also a known situation for AlaRSs from other organisms (39).

Tables 2 and 3 summarize the data collected for the new potential genes, as well as for the mt-AspRS and mt-TyrRS genes and those for the other already known human mt-aaRSs. Enzymes are sorted according to the synthetase classification, with five new specificities in each class (in boldface italic font). Table 2 collects the accession numbers, and Table 3 gives information on the genes (chromosomal location, number of exons, total length of the gene) and on

the corresponding proteins (mt-targeting probability, size of the full-length protein, predicted size of the mature enzyme, sequence identity with related human cyt-aaRS).

Each of the eight new candidate sequences shows a high mt-targeting probability as obtained by predictive programs. Interestingly, the two sets of cyt- and mt-aaRS genes are clearly distinct since they have different loci, either on different chromosomes (ArgRSs, CysRSs, GluRSs, AlaRSs, AsnRSs, and ThrRSs) or within a same chromosome (ValRSs and ProRSs). As far as sequences are concerned, mt-aaRSs show only limited identity with their cytosolic counterparts, except for the known cases of GlyRSs (12, 13) and LysRSs (19) for which it has been demonstrated that the cytosolic and mitochondrial enzymes originate from a single gene. Notice the particular high score of sequence identity of cyt- and mt-HisRSs, which has been explained by an inverted gene duplication (14, 15).

## DISCUSSION

*Human Mt-AspRS and Mt-TyrRS Are of Eubacterial Type.* Exploration of human databases revealed potential sequences for 10 mt-aaRS genes (i.e., AlaRS, ArgRS, AsnRS, AspRS, CysRS, GluRS, ProRS, ThrRS, TyrRS, and ValRS), leading thus toward the full set of mt-aaRSs. The sole exception concerns a gene for GlnRS. The new assignments are supported by (i) multiple sequences alignments with aaRS of same specificity from other organisms, (ii) the existence of typical catalytic site signature motifs, (iii) the prediction of an N-terminal mt-targeting sequence, (iv) the clear distinction between the sequences of these mt-aaRSs and their corresponding cytosolic counterparts, and (v) in the cases of mt-AspRS and mt-TyrRS, the functionality of the overexpressed and purified proteins.

Both mt-AspRS and mt-TyrRS are active enzymes, able to specifically aminoacylate *E. coli* native tRNA<sup>Asp</sup> and tRNA<sup>Tyr</sup>, as well as in vitro transcribed human mt-tRNAs.

Table 3: Characteristics of Human Mitochondrial Aminoacyl-tRNA Synthetases<sup>a</sup>

aaRSs	mt-targeting probability			genomic features			size (amino acids)		sequence comparison		
	P	MP	cleavage site	X	no. of exons	total length (nt)	full length	mature	ident %	sim %	gap %
Class I aaRSs											
<i>ArgRS</i>	0.997	0.781	after 16	VI (V)	20	75 618	578	562 (660)	28	46	8
<i>CysRS</i>	0.589	0.483	after 62	XIII (XI)	15	64 704	564	502 (748)	32	43	22
<i>GlnRS</i>				(III)				(775)			
<i>GluRS</i>	0.942	0.994	after 25	XVI (I)	9	32 980	523	498 (1440)	23	42	9
<i>IleRS</i>	0.774*	0.685	after 61	I (IX)	23	53 358	993	933 (1262)	28	44	12
<i>LeuRS</i>	0.972	0.783	<b>39 aa</b>	III (V)	20	153 073	903	<b>864</b> (1176)		not significant	
<i>MetRS</i>	0.992	0.982	<b>18 aa</b>	II (XII)	1	1 779	593	<b>575</b> (900)	24	39	10
<i>TrpRS</i>	0.995	0.735	<b>18 aa</b>	I (XIV)	6	109 446	360	<b>342</b> (471)		not significant	
<i>TyrRS</i>	0.944	0.974	after 16	XII (I)	5	8 630	477	461 (528)		not significant	
<i>ValRS</i>	0.994	0.828	after 29	VI (VI)	30	11 364	993	964 (1264)	45	58	9
Class II aaRSs											
<i>AlaRS</i>	0.973	0.998	after 68	VI (XVI)	22	13 672	985	917 (968)	45	62	3
<i>AsnRS</i>	0.984	0.974	after 14	XI (XVIII)	14	137 812	477	463 (548)	29	44	9
<i>AspRS</i>	0.816	0.998	after 47	I (II)	17	32 475	645	597 (500)	23	36	26
<i>GlyRS</i>	0.999	0.999	<b>54 aa</b>	VII (VII)	17	39 005	739	<b>685</b> (685)	100		0
<i>HisRS</i>	0.987	0.972	after 34	V (V)	13	6 900	506	472 (509)	76	87	0
<i>LysRS</i>	0.784	0.973	<b>16 aa</b>	XVI (XVI)	15	16 530	625	<b>609</b> (597)	98	98	0
<i>PheRS</i>	0.951	0.974	<b>37 aa</b>	VI ( $\alpha$ -XIX/ $\beta$ -II)	6	403 026	451	<b>414</b> ( $\alpha$ -508/ $\beta$ -589)	not significant ( $\alpha$ - or $\beta$ - chains)		
<i>ProRS</i>	0.786*	0.986	after 47	I (I)	1	1 428	475	428 (1440)	not significant		
<i>SerRS</i>	0.984	0.986	after 20	XIX (I)	16	15 127	518	498 (514)	28	49	5
<i>ThrRS</i>	0.991	0.951	after 39	I (V)	18	19 600	718	679 (712)	58	69	2

<sup>a</sup> AaRSs are sorted according to their belonging to class I (upper part) or class II (lower part) (4, 5). Mitochondrial-targeting probability was evaluated using the predictive programs Predotar (P) (or TargetP when indicated by \*) or MitoProt (MP) (for Web sites see Materials and Methods). Mt-targeting sequence cleavage sites are either predicted by the above-mentioned programs (indicated "after xx") or based on biochemical experiments (boldfaced characters). Genomic features of human mt-aaRSs include chromosomal location (X) in roman characters (those for cyt-aaRSs are recalled in parentheses), number (no.) of exons, and total gene size. Chromosomal location has been established by standard Blastn of the coding sequences against the human chromosome database, leading to hits representative of the exons. Protein sizes correspond either to full-length proteins (including the mt-targeting sequence) or to mature enzymes (in bold when experimental data are available). Protein sizes of cyt-aaRSs are recalled in parentheses. Comparison with human cyt-aaRSs indicates percentages of identity (ident), similarity (sim), and gap as calculated using standard pairwise Blast. For cyt-PheRS (composed of two subunits), comparisons have been performed with both chains. Mt-ProRS has been compared with the bifunctional cyt-GluProRS.

Catalytic activities obtained with transcripts are however less efficient (10- to 40-fold) than those obtained with native *E. coli* molecules, a situation already reported for other human mitochondrial aminoacylation systems (18, 20, 23, 36, 40, 41). This might be accounted for either (i) by a foreseen higher structural flexibility of mt-tRNAs due to their biased nucleotide composition (75% and 57% of A/U residues in mt-tRNA<sup>Asp</sup> and mt-tRNA<sup>Tyr</sup>, respectively, ref 42) or (ii) by the absence within in vitro transcripts of posttranscriptional modifications, known to have a stabilization effect on RNA folding (43) and in particular on human mt-tRNA folding (e.g., 44, 45). When tested with native *E. coli* substrates, efficiencies of the two enzymes are in a range comparable to those observed for the corresponding cognate *E. coli* enzymes tested with their native tRNA. While data are not strictly comparable due to variations in aminoacylation conditions, the efficiency of human mt-TyrRS is very close to that of *E. coli* TyrRS ( $315 \times 10^{-3} \text{ s}^{-1} \cdot \mu\text{M}^{-1}$ , ref 46), and the efficiency of mt-AspRS is  $\sim 20$ -fold lower than that of the *E. coli* enzyme (47). That some mammalian mt-aaRSs have a slower enzymatic activity than bacterial synthetases has already been documented (e.g., 18, 20). Interestingly, the *E. coli* and human mt-tRNAs are significantly different at the structural and primary sequence levels (Figure 4), which opens a number of questions as to their recognition by mt-aaRSs. Initial answers are available for both tyrosine (48) and aspartate systems (Fender et al., manuscript in preparation).

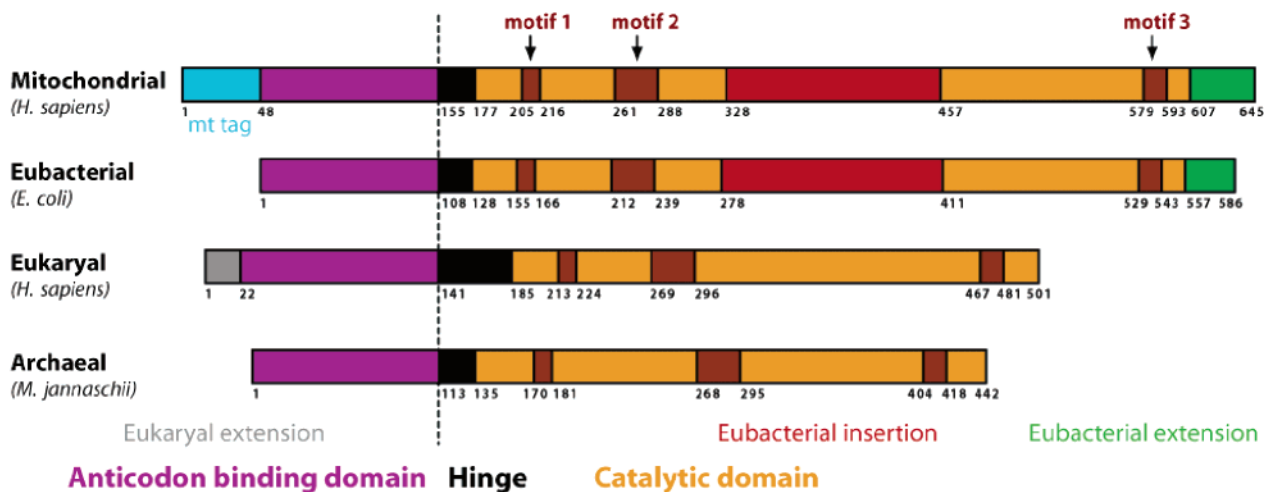
Sequence comparisons with aaRSs of same specificity show that both human mt-AspRS and mt-TyrRS fit into the expected modular organization of these families of enzymes (24, 25). The N-terminal region of mt-AspRS (Figure 5A)

corresponds to the anticodon-binding domain (aa 48–154) and its C-terminal part to the catalytic domain (aa 177–606), both connected via a hinge region. Interestingly, the mitochondrial enzyme possesses the typical eubacterial insertion (aa 328–456) and C-terminal extension (aa 607–645) domains and is thus distinguished from eukaryal and archæal AspRSs. Detailed sequence comparison within conserved domains pinpoints additional elements highlighting eubacterial features of mt-AspRS. For instance, residues T219 (corresponding to T169 in *E. coli*, located in the so-called flipping loop that closes their active sites when aspartate is bound) and Q280 (corresponding to Q231 in *E. coli*, in motif 2) are conserved solely in eubacterial AspRSs and are replaced by two serines in AspRSs from other phylae.

The modular organization of human mt-TyrRS is similar to that of other TyrRSs (Figure 5B) with an N-terminal catalytic domain (aa 32–296) followed by the C-terminal anticodon-binding region (aa 296–477). The catalytic domain presents the expected insertion domain, the so-called connective peptide CP1, here involved in dimerization and cross-species tRNA discrimination (49, 50). It also appears that the anticodon-binding region is comparable to that of eubacterial TyrRSs with similar  $\alpha$ -helical and C-terminal domains. However, on the basis of sequence comparisons, two subfamilies (represented, e.g., by *E. coli* and *T. thermophilus* TyrRSs, respectively) of eubacterial TyrRSs emerged (51). Human mt-TyrRS shares significantly higher sequence identity with the *E. coli*-type than *T. thermophilus*-type of eubacterial TyrRSs (see Figure 2). This  $\alpha$ -helical/C-terminal organization clearly distinguishes from archæal enzymes (very short and low in sequence identity at the C-terminus) and from eukaryotic TyrRSs, especially the human cytosolic



## A AspRSs



## B TyrRSs

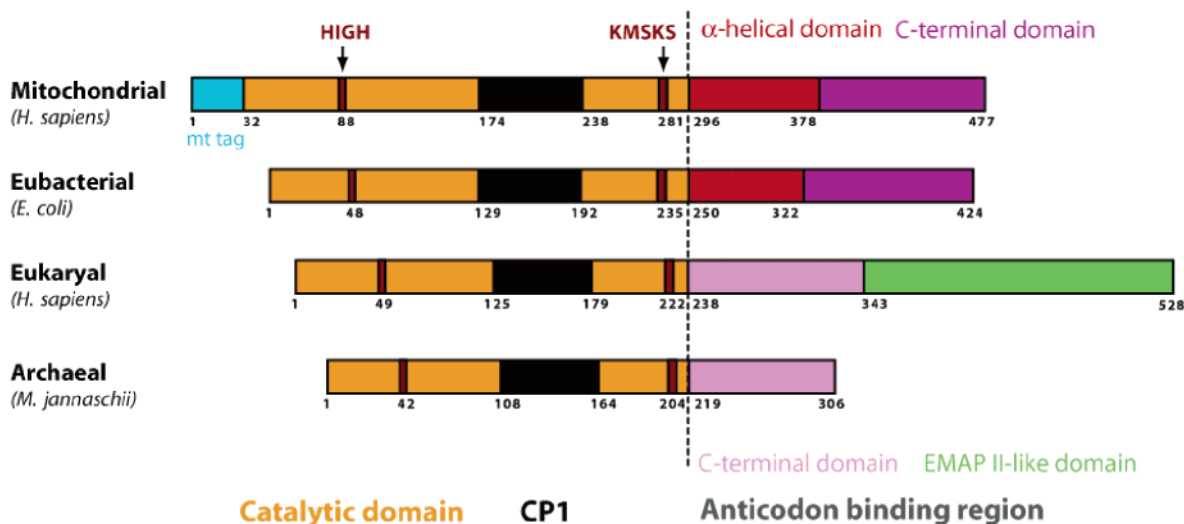


FIGURE 5: Modular organization of human mt-AspRS and mt-TyrRS compared with those from bacteria, eukarya, and archæa. The aaRSs are lined up at the junction between their anticodon-binding and catalytic domains (dotted lines). Amino acid numbering corresponds to that of the aaRSs from organisms given in parentheses. Each structural module is differently colored.

enzyme, which possesses a long EMAP II-like domain (32). This low degree in conservation for C-terminal modules is a peculiarity of TyrRSs (e.g., 25). Regarding N-terminal domains, human cyt-TyrRS presents an interleukin 8-like cytokine activity related to an ELR signature motif within the Rossmann-fold domain (52), a motif absent within mt-TyrRS (ELR is replaced by REA). Altogether, these differences (absence of EMAP II-like domain and ELR motif) suggest that no cytokine activities might be anticipated for the mitochondrial enzyme neither in its C- nor N-terminus domain. In summary, both mt-AspRS and mt-TyrRS are of eubacterial type. However, in the case of mt-TyrRS, fine inspection of the sequence of the catalytic domain shows some archæal features, supported by functional data (48).

The eubacterial character of mt-AspRS and mt-TyrRS is in line with an endosymbiotic origin of mitochondria, believed to be a remnant eubacteria engulfed by primitive eukaryotes (53, 54). In such a scenario, the present day

organelle genome would result from the progressive transfer of endosymbiont aaRS genes to the nuclear genome concomitantly with the loss of other genes. Alternatively, mt-aaRS genes could also originate from duplication of a cyt-aaRS gene (e.g., 39, 55, 56). The low sequence identity of human mt-AspRS and mt-TyrRS with the corresponding cyt-aaRSs, combined with their high sequence identities with eubacterial aaRSs and a modular organization closer to eubacterial aaRSs, favors the gene transfer hypothesis from the mitochondrial progenitor to the nucleus for these two enzymes.

*Mitochondrial and Cytosolic Human Aminoacyl-tRNA Synthetases Are Encoded by Two Sets of Genes.* Prediction of gene/protein sequences for eight additional mt-aaRSs allows now for a global view on the full set of these enzymes in human. Reconstruction of all mt-aaRS genomic sequences (established by standard Blastn of coding sequences against the human chromosome database; see Table 3) gives an

insight into their gene organization. The total length occupied by the different genes ranges from 1428 nt (ProRS) to 403 026 nt (PheRS), a variability as large as the number of exons ranking from 1 (ProRS) to 30 (ValRS). At the protein level, mature mt-aaRSs (deprived of a the mt-tag) contain (or are predicted to contain) 360–993 aa, making them 3–36% smaller than the corresponding cytosolic enzymes. As a noticeable exception, mt-AspRS is larger (596 aa) than cyt-AspRS (500 aa) due to the large eubacterial insertion domain. Interestingly, two independent coding sequences have been found for mt-GluRS and mt-ProRS, at opposite ends on a single gene for both activities in the human cytosol leading to a bifunctional GluProRS (57). Cyt-GluProRS (1440 aa) is larger than the combination of mt-GluRS (497 aa) and mt-ProRS (427 aa). Another particular situation concerns PheRSs, with the cyt-aaRS being tetrameric ( $\alpha_2\beta_2$ ), while the mt-aaRS is monomeric ( $\alpha$ ) (18). Both subunits (508 and 589 aa) of the cytosolic enzyme are larger than the single subunit of the mitochondrial enzyme (414 aa). The drastic change in mt- and cyt-PheRSs quaternary structures was also reported in yeast (58).

The striking outcome of the completed gene identification is that mitochondrial synthetases are encoded by genes different from those coding for the cyt-aaRS of same specificity. Sole exceptions concern the already reported cases of GlyRSs (12, 13) and LysRSs (19). This gene distribution is different from the situation in organisms for which sufficient information is available (reviewed, e.g., in ref 10). For example, in *S. cerevisiae* and in *Arabidopsis thaliana*, a larger number of cytosolic and mitochondrial synthetases are generated from a single gene. This is the case for yeast HisRSs (59) and ValRSs (60) and *A. thaliana* AlaRSs (61), ThrRSs, and ValRSs (62).

**Perspectives.** The identification/annotation step of the full set of human mt-aaRS genes opens also a large number of research lines at fundamental and at forensic levels. Indeed, more neurodegenerative disorders are correlated to point mutations in human mt-tRNA genes (e.g., 63–66), and defects in aminoacylation properties of these mutated tRNAs are likely targets for molecular impacts (42). Immediate access to active mt-AspRS and mt-TyrRS allows more biochemical investigation of these enzymes, including their relationships with cognate tRNAs and exploration of their aminoacylation identity elements. Experiments toward this aim are underway.

The case of mt-GlnRS remains puzzling. We were unable to detect a candidate gene for this enzyme. Among possible explanations, it can be proposed either that the sequence of human mt-GlnRS has evolved so much that it has become unrelated to any of the known GlnRSs from other organisms or that its function is fulfilled by a mt-addressed version of the cyt-GlnRS. So far, we were unable to detect sequences that may unravel these two hypotheses. Another alternative, in line with previous reports suggesting that this enzyme does not exist (67, 68), would be that synthesis of mt-Gln-tRNA<sup>Gln</sup> occurs via an indirect pathway involving two steps, misacylation of tRNA<sup>Gln</sup> by GluRS yielding Glu-tRNA<sup>Gln</sup> followed by Glu-amidation (69). This hypothesis is supported by the existence of a gene homologous to glutaminyl-amidotransferases in the human genome (accession number NP\_004555). However, further investigations are needed to

decipher the glutaminylation pathway in human mitochondria.

From an evolutionary point of view, the present work further confirms and extends the view that human mt- and cyt-aaRSs are coded by two different sets of genes. However, in humans, there are two exceptions as described above. In other organisms, this difference can be greater like in yeast and *A. thaliana*. An outcoming question is to know whether a situation could occur in nature where the two sets of genes are completely independent. Inspection of the newly sequenced genomes should shed light on this problem.

## ACKNOWLEDGMENT

We thank Magali Frugier and Bernard Lorber for fruitful advice, Franck Martin and Daniel Kern for gift of materials, and Olivier Poch and Gilbert Eriani for helpful discussions.

## REFERENCES

- Martinis, S. A., Plateau, P., Cavarelli, J., and Florentz, C. (1999) Aminoacyl-tRNA synthetases: a new image for a classical family, *Biochimie* 81, 683–700.
- Francklyn, C., Perona, J. J., Pütz, J., and Hou, Y.-M. (2002) Aminoacyl-tRNA synthetases: Versatile players in the changing theater of translation, *RNA* 8, 1363–1372.
- Ibba, M., Francklyn, C., and Cusack, S. (2004) *Aminoacyl-tRNA Synthetases*, Landes Biosciences, Georgetown, TX.
- Eriani, G., Delarue, M., Poch, O., Gangloff, J., and Moras, D. (1990) Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs, *Nature* 347, 203–206.
- Cusack, S., Berthet-Colominas, C., Härtlein, M., Nassar, N., and Leberman, R. (1990) A second class of synthetase structure revealed by X-ray analysis of *Escherichia coli* seryl-tRNA synthetase, *Nature* 347, 249–255.
- Hendrickson, T. L., and Schimmel, P. (2003) in *Translation Mechanisms* (Lapointe, J., and Brakier-Gingras, L., Eds.) pp 34–64, Landes Bioscience, Georgetown, TX.
- McClain, W. H. (1993) Rules that govern tRNA identity in protein synthesis, *J. Mol. Biol.* 234, 257–280.
- Giegé, R., Sissler, M., and Florentz, C. (1998) Universal rules and idiosyncratic features in tRNA identity, *Nucleic Acids Res.* 26, 5017–5035.
- Beuning, P. J., and Musier-Forsyth, K. (1999) Transfer RNA recognition by aminoacyl-tRNA synthetases, *Biopolymers* 52, 1–28.
- Sissler, M., Pütz, J., Fasiolo, F., and Florentz, C. (2004) Mitochondrial aminoacyl-tRNA synthetases, in *Aminoacyl-tRNA synthetases* (Ibba, M., Francklyn, C., and Cusack, S., Eds.) Chapter 28, Landes Biosciences, Georgetown, TX.
- Chomyn, A., Cleeter, M. W., Ragan, C. I., Riley, M., Doolittle, R. F., and Attardi, G. (1986) URF6, last unidentified reading frame of human mtDNA, codes for an NADH dehydrogenase subunit, *Science* 234, 614–618.
- Shiba, K., Schimmel, P., Motegi, H., and Noda, T. (1994) Human glycyl-tRNA synthetase. Wide divergence of primary structure from bacterial counterpart and species-specific aminoacylation, *J. Biol. Chem.* 269, 30049–30055.
- Mudge, S. J., Williams, J. H., Eyre, H. J., Sutherland, G. R., Cowan, P. J., and Power, D. A. (1998) Complex organisation of the 5'-end of the human glycine tRNA synthetase gene, *Gene* 209, 45–50.
- O'Hanlon, T. P., Raben, N., and Miller, F. W. (1995) A novel gene oriented in a head-to-head configuration with the human histidyl-tRNA synthetase (HRS) gene encodes an mRNA that predicts a polypeptide homologous to HRS, *Biochem. Biophys. Res. Commun.* 210, 556–566.
- O'Hanlon, T. P., and Miller, F. W. (2002) Genomic organization, transcriptional mapping, and evolutionary implications of the human bi-directional histidyl-tRNA synthetase locus (HARS/HARSL), *Biochem. Biophys. Res. Commun.* 294, 609–614.
- Shiba, K., Suzuki, N., Shigesada, K., Namba, Y., Schimmel, P., and Noda, T. (1994) Human cytoplasmic isoleucyl-tRNA synthetase: Selective divergence of the anticodon-binding domain

- and acquisition of a new structural unit, *Proc. Natl. Acad. Sci. U.S.A.* 91, 7435–7439.
17. Kelley, S., Steinberg, S., and Schimmel, P. (2000) Functional defects of pathogenic human mitochondrial tRNAs related to structural fragility, *Nat. Struct. Biol.* 7, 862–865.
  18. Bullard, J., Cai, Y.-C., Demeler, B., and Spremulli, L. (1999) Expression and characterization of a human mitochondrial phenylalanyl-tRNA synthetase, *J. Mol. Biol.* 288, 567–577.
  19. Tolkunova, E., Park, H., Xia, J., King, M. P., and Davidson, E. (2000) The human lysyl-tRNA synthetase gene encodes both the cytoplasmic and mitochondrial enzymes by means of an unusual splicing of the primary transcript, *J. Biol. Chem.* 275, 35063–35069.
  20. Bullard, J., Cai, Y.-C., and Spremulli, L. (2000) Expression and characterization of the human mitochondrial leucyl-tRNA synthetase, *Biochim. Biophys. Acta* 1490, 245–258.
  21. Jørgensen, R., Søgarrd, M. M., Rossing, A. B., Martensen, P. M., and Justesen, J. (2000) Identification and characterization of human mitochondrial tryptophanyl-tRNA synthetase, *J. Biol. Chem.* 275, 16820–16826.
  22. Yokogawa, T., Shimada, N., Takeuchi, N., Benkowski, L., Suzuki, T., Omori, A., Ueda, T., Nishikawa, K., Spremulli, L. L., and Watanabe, K. (2000) Characterization and tRNA recognition of mammalian mitochondrial seryl-tRNA synthetase, *J. Biol. Chem.* 275, 19913–19920.
  23. Spencer, A., Heck, A., Takeuchi, N., Watanabe, K., and Spremulli, L. (2004) Characterization of the human mitochondrial methionyl-tRNA synthetase, *Biochemistry* 43, 9743–9754.
  24. Giegé, R., and Rees, B. (2004) Aspartyl-tRNA synthetases, in *Aminoacyl-tRNA Synthetases* (Ibba, M., Francklyn, C., and Cusack, S., Eds.) Chapter 22, Landes Bioscience, Georgetown, TX.
  25. Bedouelle, H. (2004) Tyrosyl-tRNA synthetases, in *Aminoacyl-tRNA Synthetases* (Ibba, M., Francklyn, C., and Cusack, S., Eds.) Chapter 13, Landes Biosciences, Georgetown, TX.
  26. Frugier, M., Moulinier, L., and Giegé, R. (2000) A domain in the N-terminal extension of class IIb eukaryotic aminoacyl-tRNA synthetases is important for tRNA binding, *EMBO J.* 19, 2371–2380.
  27. Mirande, M., Gache, Y., Le, C. D., and Waller, J. P. (1982) Seven mammalian aminoacyl-tRNA synthetases co-purified as high molecular weight entities are associated within the same complex, *EMBO J.* 1, 733–736.
  28. Agou, F., and Mirande, M. (1997) Aspartyl-tRNA synthetase from rat: in vitro functional analysis of its assembly into the multi-synthetase complex, *Eur. J. Biochem.* 243, 259–267.
  29. Lee, C. P., and RajBhandary, U. L. (1991) Mutants of *Escherichia coli* initiator tRNA which suppress amber codons in *Saccharomyces cerevisiae* and are aminoacylated with tyrosine by yeast extracts, *Proc. Natl. Acad. Sci. U.S.A.* 88, 11378–11382.
  30. Yaremchuk, A., Krikliviy, I., Tukalo, M., and Cusack, S. (2002) Class I tyrosyl-tRNA synthetase has a class II mode of cognate tRNA recognition, *EMBO J.* 21, 3829–3840.
  31. Steer, B. A., and Schimmel, P. (1999) Major anticodon-binding region missing from an archaeobacterial tRNA synthetase, *J. Biol. Chem.* 274, 35601–35606.
  32. Wakasugi, K., and Schimmel, P. (1999) Two distinct cytokines released from a human aminoacyl-tRNA synthetase, *Science* 284, 147–150.
  33. Becker, H. D., Giegé, R., and Kern, D. (1996) Identity of prokaryotic and eukaryotic tRNA<sup>Asp</sup> for aminoacylation by aspartyl-tRNA synthetase from *Thermus thermophilus*, *Biochemistry* 35, 7447–7458.
  34. Fechter, P., Rudinger, J., Giegé, R., and Théobald-Dietrich, A. (1998) Ribozyme processed tRNA transcripts with unfriendly internal promoter for T7 RNA polymerase: production and activity, *FEBS Lett.* 436, 99–103.
  35. Sissler, M., Helm, M., Frugier, M., Giegé, R., and Florentz, C. (2004) Aminoacylation properties of pathology-related variants of human mitochondrial tRNA<sup>Lys</sup> variants, *RNA* 10, 841–853.
  36. Sohm, B., Frugier, M., Brulé, H., Olszak, K., Przykorska, A., and Florentz, C. (2003) Towards understanding human mitochondrial leucine aminoacylation identity, *J. Mol. Biol.* 328, 995–1010.
  37. Helm, M., Brulé, H., Friede, D., Giegé, R., Pütz, J., and Florentz, C. (2000) Search for characteristic structural features of mammalian mitochondrial tRNAs, *RNA* 6, 1356–1379.
  38. Delagoutte, B., Moras, D., and Cavarelli, J. (2000) tRNA aminoacylation by arginyl-tRNA synthetase: induced conformations during substrates binding, *EMBO J.* 19, 5599–5610.
  39. Chihade, J. W., Brown, J. R., Schimmel, P., and Ribas de Pouplana, L. (2000) Origin of mitochondria in relation to evolutionary history of eukaryotic alanyl-tRNA synthetase, *Proc. Natl. Acad. Sci. U.S.A.* 97, 12153–12157.
  40. Degoul, F., Brulé, H., Cepanec, C., Helm, M., Marsac, C., Leroux, J.-P., Giegé, R., and Florentz, C. (1998) Isoleucylation properties of native human mitochondrial tRNA<sup>Ile</sup> and tRNA<sup>Ile</sup> transcripts. Implications for cardiomyopathy-related point mutations (4269, 4317) in the tRNA<sup>Ile</sup> gene, *Hum. Mol. Genet.* 7, 347–354.
  41. Park, H., Davidson, E., and King, M. P. (2003) The pathogenic A3243G mutation in human mitochondrial tRNA<sup>Leu(UUR)</sup> decreases the efficiency of aminoacylation, *Biochemistry* 4, 958–964.
  42. Florentz, C., Sohm, B., Tryoen-Tóth, P., Pütz, J., and Sissler, M. (2003) Human mitochondrial tRNAs in health and disease, *Cell. Mol. Life Sci.* 60, 1356–1375.
  43. Hall, K. B., Sampson, J. R., Uhlenbeck, O. C., and Redfield, A. G. (1989) Structure of an unmodified tRNA molecule, *Biochemistry* 28, 5794–5801.
  44. Helm, M., Brulé, H., Degoul, F., Cepanec, C., Leroux, J.-P., Giegé, R., and Florentz, C. (1998) The presence of modified nucleotides is required for cloverleaf folding of a human mitochondrial tRNA, *Nucleic Acids Res.* 26, 1636–1643.
  45. Sohm, B., Sissler, M., Park, H., King, M. P., and Florentz, C. (2004) Recognition of human mitochondrial tRNA<sup>Leu(UUR)</sup> by its cognate leucyl-tRNA synthetase, *J. Mol. Biol.* 339, 17–29.
  46. Bedouelle, H., and Winter, G. (1986) A model of synthetase/transfer RNA interaction as deduced by protein engineering, *Nature* 320, 371–373.
  47. Moulinier, L., Eiler, S., Eriani, G., Gangloff, J., Thierry, J.-C., Gabriel, K., McClain, W. H., and Moras, D. (2001) The structure of an AspRS-tRNA<sup>Asp</sup> complex reveals a tRNA-dependent control mechanism, *EMBO J.* 20, 5290–5301.
  48. Bonnefond, L., Frugier, M., Giegé, R., and Rudinger-Thirion J. (2005) Human mitochondrial TyrRS disobeys the tyrosine identity rules, *RNA*, in press.
  49. Wakasugi, K., Quinn, C. L., Tao, N., and Schimmel, P. (1998) Genetic code in evolution: switching species-specific aminoacylation with a peptide transplant, *EMBO J.* 17, 297–305.
  50. Kobayashi, T., Nureki, O., Ishitani, R., Yaremchuk, A., Tukalo, M., Cusack, S., Sakamoto, K., and Yokoyama, S. (2003) Structural basis for orthogonal tRNA specificities of tyrosyl-tRNA synthetases for genetic code expansion, *Nat. Struct. Biol.* 10, 425–432.
  51. Salazar, J. C., Zuniga, R., Lefimil, C., Söll, D., and Orellana, O. (2001) Conserved amino acids near the carboxy terminus of bacterial tyrosyl-tRNA synthetase are involved in ntRA and Tyr-AMP binding, *FEBS Lett.* 491, 257–260.
  52. Wakasugi, K., and Schimmel, P. (1999) Highly differentiated motifs responsible for two cytokine activities of a spilt human tRNA synthetase, *J. Biol. Chem.* 274, 23155–23159.
  53. Cavalier-Smith, T. (1987) The simultaneous symbiotic origin of mitochondria, chloroplasts, and microbodies, *Ann. N.Y. Acad. Sci.* 503, 55–71.
  54. Gray, M. W. (2000) Mitochondrial genes on the move, *Nature* 408, 302–305.
  55. Ribas de Pouplana, L., and Schimmel, P. (2000) A view into the origin of life: aminoacyl-tRNA synthetases, *Cell. Mol. Life Sci.* 57, 865–870.
  56. Adams, K. L., and Palmer, J. D. (2003) Evolution of mitochondrial gene content: gene loss and transfer to the nucleus, *Mol. Phylogenet. Evol.* 29, 380–395.
  57. Kaiser, E., Hu, B., Becher, S., Eberhard, D., Schray, B., Baack, M., Hameister, H., and Knippers, R. (1994) The human EPRS locus (formerly the QARS locus): a gene encoding a class I and a class II aminoacyl-tRNA synthetase, *Genomics* 19, 280–290.
  58. Sanni, A., Walter, P., Boulanger, Y., Ebel, J.-P., and Fasiolo, F. (1991) Evolution of aminoacyl-tRNA synthetase quaternary structure and activity: *Saccharomyces cerevisiae* mitochondrial phenylalanyl-tRNA synthetase, *Proc. Natl. Acad. Sci. U.S.A.* 88, 8387–8391.
  59. Natsoulis, G., Hilger, F., and Fink, G. R. (1986) The HTS1 gene encodes both the cytoplasmic and mitochondrial histidine tRNA synthetase of *S. cerevisiae*, *Cell* 46, 235–243.
  60. Tettelin, H., Agostoni Carbone, M. L., Albermann, K., Albers, M., Arroyo, J., Backes, U., Barreiros, T., Bertani, I., Bjourson, A. J., Bruckner, M., Bruschi, C. V., Carignani, G., Castagnoli, L., Cerdan, E., Clemente, M. L., Coblenz, A., Cogliervina, M., Coissac, E., Defoor, E., Del Bino, S., Delius, H., Delneri, D., de

- Wergifosse, P., Dujon, B., Kleine, K., et al. (1997) The nucleotide sequence of *Saccharomyces cerevisiae* chromosome VII, *Nature* 387, 81–84.
61. Mireau, H., Lancelin, D., and Small, I. D. (1996) The same *Arabidopsis* gene encodes both cytosolic and mitochondrial alanyl-tRNA synthetases, *Plant Cell* 8, 1027–1039.
62. Souciet, G., Menand, B., Ovesna, J., Cosset, A., Dietrich, A., and Wintz, H. (1999) Characterization of two bifunctional *Arabidopsis thaliana* genes coding for mitochondrial and cytosolic forms of valyl-tRNA synthetase and threonyl-tRNA synthetase by alternative use of two in-frame AUGs, *Eur. J. Biochem.* 266, 848–854.
63. Schon, E. A., Bonilla, E., and DiMauro, S. (1997) Mitochondrial DNA mutations and pathogenesis, *J. Bioenerg. Biomembr.* 29, 131–149.
64. Wallace, D. C. (1999) Mitochondrial diseases in man and mouse, *Science* 283, 1482–1488.
65. Schon, E. (2000) Mitochondrial genetics and disease, *Trends Biochem. Sci.* 25, 555–560.
66. DiMauro, S., and Andreu, A. (2000) Mutations in mtDNA: are we scraping the bottom of the barrel? *Brain Pathol.* 10, 431–441.
67. Schön, A., Kannangara, G., Gough, S., and Söll, D. (1988) Protein biosynthesis in organelles requires misaminoacylation of tRNA, *Nature* 331, 187–190.
68. Schön, A., and Söll, D. (1988) tRNA specificity of a mischarging aminoacyl-tRNA synthetase: Glutamyl-tRNA synthetase from barley chloroplasts, *FEBS Lett.* 228, 241–244.
69. Ibba, M., and Söll, D. (2000) Aminoacyl-tRNA synthesis, *Annu. Rev. Biochem.* 69, 617–650.
70. Poirrot, O., Suhre, K., Abergel, C., O'Toole, E., and Notredame, C. (2004) 3DCoffee@igs: a web server for combining sequences and structures into multiple sequence alignment, *Nucleic Acids Res.* 32, W37–W40.
71. Rost, B., Yachdav, G., and Liu, J. (2004) The PredictProtein server, *Nucleic Acids Res.* 32, 321–326.
72. Gouet, P., Courcelle, E., Stuart, D. I., and Metz, F. (1999) ESPript: analysis of multiple sequence alignments in PostScript, *Bioinformatics* 15, 305–308.
73. Anderson, S., Banker, A. T., Barrel, B. G., de Bruijn, M. H. L., Coulson, A. R., Drouin, J., Eperon, J. C., Nierlich, D. P., Roe, B. A., Sanger, F., et al. (1981) Sequence and organization of the human mitochondrial genome, *Nature* 290, 457–465.
74. Sprinzl, M., Horn, C., Brown, M., Ioudovitch, A., and Steinberg, S. (1998) Compilation of tRNA sequences and sequences of tRNA genes, *Nucleic Acids Res.* 26, 148–153.

BI047527Z



## **B. Un jeu d'éléments d'identité restreint**

### **1. Perte de spécificité**

Le rôle de l'élément d'identité N1-N72, central pour la spécificité des ARNt<sup>Tyr</sup> eubactériens et eucaryotiques, a été le premier élément que nous avons cherché à élucider pour le système mitochondrial humain. Les résultats obtenus ont fait l'objet d'une publication sous la forme d'une lettre à l'éditeur dans la revue "RNA" (en 2005) et ont été présentés par affiches au "Forum des Jeunes Chercheurs" (Nantes, Octobre 2005) ainsi qu'au "21<sup>st</sup> International tRNA Workshop" (Bangalore, Inde, Décembre 2006).

Article n°3

### **La TyrRS mitochondriale humaine enfreint les règles de l'identité tyrosine**

Luc Bonnefond, Magali Frugier, Richard Giegé et Joëlle Rudinger-Thirion

**2005**

RNA, **11** : 558-562

La tyrosyl-ARNt synthétase mitochondriale humaine (mt-TyrRS) présente des éléments de séquence caractéristiques des TyrRS eubactériennes mais aussi d'archaea, en particulier au niveau de la région reconnaissant l'élément d'identité N1-N72 de l'ARNt<sup>Tyr</sup>. Ce constat suggère que la TyrRS mitochondriale humaine a perdu sa capacité à discriminer entre la paire G1-C72 typique des bactéries/mitochondries et la paire inverse C1-G72 présente chez les ARNt<sup>Tyr</sup> d'archaea/eucaryotes. Cette hypothèse a été validée par l'analyse de la capacité de la mt-TyrRS à aminoacyler des molécules d'ARNt<sup>Tyr</sup> sauvages ou mutées. Les résultats montrent que la TyrRS mitochondriale aminoacyle aussi efficacement son ARNt<sup>Tyr</sup> homologue avec une paire G1-C72 que sa version mutée en une paire C1-G72. Cette mt-TyrRS est ainsi le premier exemple d'une TyrRS dépourvue de spécificité vis-à-vis de la première paire de bases N1-N72 et ne respectant pas les règles d'identité des systèmes d'aminocyclation spécifiques de la tyrosine. La comparaison des séquences de diverses TyrRS mitochondriales suggère que cette particularité fonctionnelle est conservée chez les mt-TyrRS de vertébrés.

# Human mitochondrial TyrRS disobeys the tyrosine identity rules

LUC BONNEFOND, MAGALI FRUGIER, RICHARD GIEGÉ, and JOËLLE RUDINGER-THIRION

Département Mécanismes et Macromolécules de la Synthèse Protéique et Cristallogénèse, Unité Propre de Recherche 9002 du Centre National de la Recherche Scientifique, Institut de Biologie Moléculaire et Cellulaire, F-67084 Strasbourg, France

## ABSTRACT

Human tyrosyl-tRNA synthetase from mitochondria (mt-TyrRS) presents dual sequence features characteristic of eubacterial and archaeal TyrRSs, especially in the region containing amino acids recognizing the N1-N72 tyrosine identity pair. This would imply that human mt-TyrRS has lost the capacity to discriminate between the G1-C72 pair typical of eubacterial and mitochondrial tRNA<sup>Tyr</sup> and the reverse pair C1-G72 present in archaeal and eukaryal tRNA<sup>Tyr</sup>. This expectation was verified by a functional analysis of wild-type or mutated tRNA<sup>Tyr</sup> molecules, showing that mt-TyrRS aminoacylates with similar catalytic efficiency its cognate tRNA<sup>Tyr</sup> with G1-C72 and its mutated version with C1-G72. This provides the first example of a TyrRS lacking specificity toward N1-N72 and thus of a TyrRS disobeying the identity rules. Sequence comparisons of mt-TyrRSs across phylogeny suggest that the functional behavior of the human mt-TyrRS is conserved among all vertebrate mt-TyrRSs.

**Keywords:** tRNA identity; cross-species aminoacylation; phylogeny; tRNA acceptor stem; tyrosylation

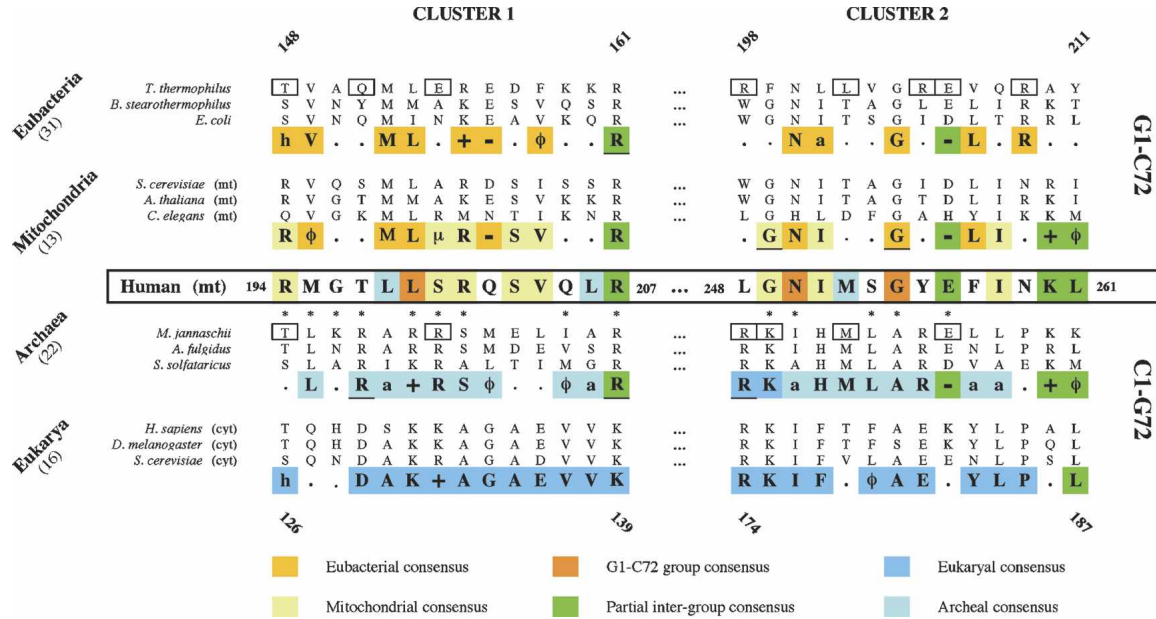
Tyrosyl-tRNA synthetase (TyrRS) is distinguished from other aminoacyl-tRNA synthetases by its ambiguous assignment to class I of synthetases (Bedouelle 2004). Indeed, this homodimeric enzyme with typical class I motifs binds its homologous tRNA<sup>Tyr</sup> across its interface (Bedouelle and Winter 1986) via the major groove side of the amino acid acceptor stem, as occurs for class II synthetases (Lee and RajBhandary 1991; Bedouelle et al. 1993; Fechter et al. 2000; Yaremchuk et al. 2002). Furthermore, tyrosylation reactions are peculiar in that no cross-aminoacylation between eubacterial tRNA<sup>Tyr</sup> and eukaryal TyrRS is possible and vice-versa (Kleeman et al. 1997; Wakasugi et al. 1998). This strong and unusual phylogenetic barrier mainly relies on the properties of the N1-N72 base pair at the top of the acceptor stem that was assigned as the major tyrosine identity element in eukarya (C1-G72) (Fechter et al. 2000) and a minor element in eubacteria (G1-C72) (Quinn et al. 1995). Sequence alignments reveal strict conservation of the G1-C72 pair in all eubacterial tRNA<sup>Tyr</sup> species and its replacement by the reverse C1-G72 pair in eukaryal and archaeal tRNA<sup>Tyr</sup> molecules, an uncommon pair at this position in other tRNAs (Sprinzl and Vassilenko 2005). Recently, the crystallographic structures of *Thermus thermo-*

*philus* and *Methanococcus jannaschii* tRNA<sup>Tyr</sup>/TyrRS complexes brought to light the specific interactions established between the N1-N72 pair and TyrRS (Yaremchuk et al. 2002; Giegé 2003; Kobayashi et al. 2003). The differences seen in these interactions provide a structural support accounting for the absence of cross-tyrosylation. In this context, the correlation between two specific sequence clusters in the catalytic domain of TyrRSs and the N1-N72 identity pair can be recalled (Nair et al. 1997). Indeed, cluster 1, within the connective peptide domain, contains the main amino acids contacting the bases of the N1-N72 identity pair (Yaremchuk et al. 2002; Kobayashi et al. 2003). The functional role of these amino acids (Fig. 1) was already supported by peptide transplantation experiments enabling an identity switch in heterologous tyrosylation systems (Wakasugi et al. 1998). Likewise, cluster 2 is involved in N1 recognition (Yaremchuk et al. 2002; Kobayashi et al. 2003).

On the other hand, beside a few studies on yeast mt-TyrRSs involved in splicing events (Akins and Lambowitz 1987; Kamper et al. 1992), nothing is known concerning other mt-TyrRSs, in particular in relation to tyrosine identity. In the frame of a global approach to study human mitochondrial synthetases, the human mt-TyrRS has been purified (Bonnetfond et al. 2005). Sequence comparisons of this enzyme with other TyrRSs revealed a eubacterial character as could be anticipated from the close evolutionary relation between mitochondria and eubacteria (Gray et al. 2001). Surprisingly, human mt-TyrRS presents also archaeal

**Reprint requests to:** Richard Giegé, UPR 9002 du CNRS-IBMC, 15 rue René Descartes, F-67084 Strasbourg, France; e-mail: r.giege@ibmc.u-strasbg.fr; fax: 33-(0)3-88-60-22-18

Article and publication are at <http://www.rnajournal.org/cgi/doi/10.1261/rna.7246805>.

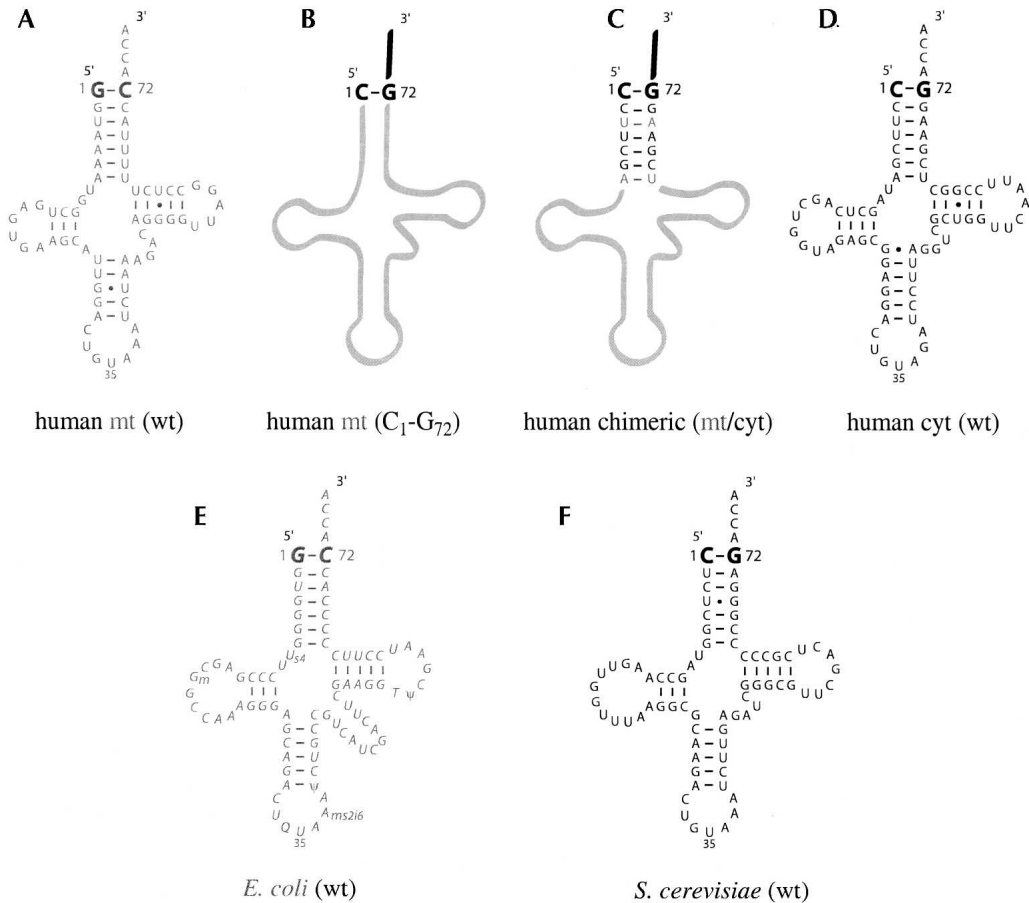


**FIGURE 1.** Sequence comparison of the clusters in the catalytic domains of TyrRSs in the vicinity to identity elements in tRNA<sup>Tyr</sup>. The figure emphasizes the clusters of human mt-TyrRS and compares their sequences to those of TyrRSs from four phylogenetic groups ranked according to the nature of the N1-N72 base pair. For each group (eubacteria, mitochondria—but without vertebrate mitochondria—archaea, eukarya) three sequences are explicitly given. Among mitochondrial TyrRSs, those of *Arabidopsis thaliana* and *Caenorhabditis elegans* were annotated for the purpose of this work. Alignments were done with Tcoffee (Poirot et al. 2003) or 3DCoffee (Poirot et al. 2004) to take into account the known TyrRS crystallographic structures (*T. thermophilus*, Yaremchuk et al. 2002; *M. jannaschii*, Kobayashi et al. 2003; *Bacillus stearothermophilus*, Brick et al. 1989; *Staphylococcus aureus*, Qiu et al. 2001; and *Homo sapiens* mini-TyrRS, Yang et al. 2002). Numberings flanking the alignments correspond to those of TyrRS from *T. thermophilus* for eubacterial/mitochondrial TyrRSs and from *M. jannaschii* for archaeal/eukaryal TyrRSs; numbering of human mt-TyrRS is also given. The asterisks below the human mitochondrial sequence indicate strict conservation of amino acids in vertebrate mt-TyrRSs (see the text). Consensus sequences were established with up to 31 known or predicted sequences (exact numbers are indicated in parentheses); the displayed consensus residues are present in >70% of the analyzed sequences. Underlined amino acids are strictly conserved; amino acids with the same properties are depicted by (ϕ) for hydrophobic, (+) positively charged, (−) negatively charged, (a) aliphatic, (μ) small, and (h) with hydroxyl group. The consensus within the G1-C72 and C1-G72 groups and within the three domains of life or mitochondria are emphasized by colors (notice that the conserved amino acids crossing the G1-C72/C1-G72 barrier are in green). Squared residues are those found in direct contact with bases of either *T. thermophilus* tRNA<sup>Tyr</sup> (Yaremchuk et al. 2002) or *M. jannaschii* tRNA<sup>Tyr</sup> (Kobayashi et al. 2003).

sequence features in the two aforementioned clusters (Fig. 1). When comparing the residues in human mt-TyrRS homologous to those interacting directly with the acceptor stem of tRNA<sup>Tyr</sup>, as seen in the crystal structures (Yaremchuk et al. 2002; Kobayashi et al. 2003), it appears that four among them cannot be linked to any phylogenetic group of TyrRSs (T197, L248, Y255, and N259), three are of mitochondrial type (R194, S200, and G249), M252 is of archaeal type, and E256 is conserved in almost all TyrRSs. It was previously suggested and experimentally verified that the sequence of these clusters relies on the specificity of the synthetase toward tRNA (Quinn et al. 1995). In the present case, the mosaic character of the human mt-TyrRS in the two clusters implies a different recognition of the tRNA than that found typically in eubacteria or eukarya. In other words, human mt-TyrRS would tyrosylate tRNA<sup>Tyr</sup> species disregarding the nature of their N1-N72 base pair. To verify this assumption, a series of tRNA<sup>Tyr</sup> molecules from different organisms, either wild type or mutated, were tested for their tyrosylation abilities (Fig. 2).

It is known that human mt-TyrRS charges not only its

cognate mitochondrial tRNA<sup>Tyr</sup> transcript but also, and even more efficiently, *Escherichia coli* tRNA<sup>Tyr</sup>, both molecules presenting a G1-C72 pair at the top of their acceptor stems (Bonfond et al. 2005; Fig. 2A,E; Table 1). The importance of this pair for eubacterial TyrRS recognition has been known since the 1970s (Celis et al. 1973) and was strengthened by studies on *E. coli* microhelices showing that a transversion of G1-C72 into C1-G72 reduces the efficiency of tyrosylation by the homologous enzyme (Quinn et al. 1995). Because of the mosaic character of human mt-TyrRS, a mutant of its cognate tRNA<sup>Tyr</sup> (Fig. 2B) with the reverse eukaryal/archaeal type identity version (C1-G72) should also present tyrosylation capacity. This was verified by a kinetic analysis showing that permutation of the N1-N72 pair does not affect the overall tyrosylation efficiency (as defined by the  $k_{cat}/K_M$  ratio) of these tRNAs (Table 1). However, the individual kinetic parameters are modified, with reduced  $K_M$  and  $k_{cat}$  values, indicating an increased affinity of the mutant for mt-TyrRS compensated by a reduced turnover of the enzyme. This functional behavior demonstrates the noninvolvement of the N1-N72 pair in



**FIGURE 2.** Sequence of wild-type and mutated tRNA<sup>Tyr</sup> molecules. The nature of the N1-N72 base pair is emphasized in each molecule. Wild-type human mt-tRNA<sup>Tyr</sup> is depicted in A, its variant with the reversed C1-G72 pair in B, chimeric mt/cyt-tRNA<sup>Tyr</sup> with cyt-tRNA<sup>Tyr</sup> acceptor stem in C, and wild-type human cyt-tRNA<sup>Tyr</sup> in D. The sequence of native tRNA<sup>Tyr</sup> from *E. coli* (E) is displayed in italics. Mitochondrial and *E. coli* sequences are in gray whereas eukaryal (cytosolic) sequences are in black. Sequence data are taken from Sprinzl and Vassilenko (2005). All transcripts were produced in vitro using the “transzyme” method and purified as described in Fechter et al. (1998) whereas native tRNA<sup>Tyr</sup> from *E. coli* (E) was from Subriden.

mitochondrial tyrosine identity in human, although the altered values of  $K_M$  and  $k_{cat}$  suggest that N1 and/or N72 play a role in the catalytic mechanism of tyrosylation. To our knowledge, this is the first case where identity rules are disobeyed with a determinant having lost its universal function. From the point of view of evolution, this conclusion may appear surprising since identity rules are ancient. To rationalize this contradiction, one can suggest that the tyrosine system in human mitochondria was under a subtle evolutionary process where mutations in the synthetase relaxed its specificity by removing or altering specific contacts with the primordial G1-C72 identity pair. This process should find appropriate conditions within mitochondria that are niches with a reduced set of tRNAs and a simplified and more permissive macromolecular environment. This view finds support in the overall different tyrosylation mechanism of tRNAs with G1-C72 or G1-C72 pairs.

In all aminoacylation systems studied so far, part of the acceptor branch was always found involved in identity

(Giegé et al. 1998; Beuning and Musier-Forsyth 1999). Therefore we first verified whether other elements of the acceptor stem participate in tyrosine identity. For that we constructed a chimeric molecule derived from wild-type human mt-tRNA<sup>Tyr</sup> with the whole stem replaced by that of human cytosolic (cyt) tRNA<sup>Tyr</sup>. This variant has a C1-G72 pair and only two (out of seven) base pairs in common with human mt-tRNA<sup>Tyr</sup> (U3-A70 and A7-U65) (Fig. 2C). Based on the crystal structure of the *T. thermophilus* TyrRS/tRNA<sup>Tyr</sup> complex (Yaremchuk et al. 2002), these two pairs would a priori not interact with mt-TyrRS, and therefore not participate in tyrosine identity. This is also the case for the other four remaining pairs, since the chimerical molecule is fully active in tyrosylation and, remarkably, displays kinetic parameters similar to those of wild-type mt-tRNA<sup>Tyr</sup> (Table 1). Thus, it can be concluded for an absence of tyrosine identity determinants in the double-stranded moiety of the acceptor branch in human mt-tRNA<sup>Tyr</sup>. This conclusion is supported by the efficient tyrosylation by hu-



**TABLE 1.** Tyrosylation by human mt-TyrRS of wild-type tRNA<sup>Tyr</sup> molecules or their variants with mutations in the acceptor stem

tRNA <sup>Tyr</sup>		$K_M$ ( $\mu\text{M}$ )	$k_{\text{cat}}$ ( $10^{-3} \cdot \text{s}^{-1}$ )	$k_{\text{cat}}/K_M$ ( $10^{-3} \cdot \text{s}^{-1} \cdot \mu\text{M}^{-1}$ )	$L^a$ (x-fold)
[G1-C72] group					
human mt <sup>b</sup>	(wt)	4.8	46	9.6	1
<i>E. coli</i> <sup>b</sup>	(wt)	0.52	210	400	0.024
human mt (A73)	(var)	24	0.054	0.002	4800
[C1-G72] group					
human mt (C1-G72)	(var)	0.48	3.4	7.1	1.3
human chimeric (mt/cyt)	(var)	2.2	30	13.6	0.7
human cyt	(wt)	48	100	2.1	4.6
<i>S. cerevisiae</i>	(wt)	1.3	1.7	1.3	7.4

tRNAs are ranked in two groups according to the nature of the N1-N72 pair. The C-terminal His-tagged human mt-TyrRS was cloned, expressed, and purified as described elsewhere (Bonnefond et al. 2005). Tyrosylation assays were performed at 37°C in 50 mM HEPES-NaOH (pH 7.6) buffer containing 25 mM KCl, 12 mM MgCl<sub>2</sub>, 2.5 mM ATP, 0.2 mg/mL bovine serum albumin, 1 mM spermine, and 10  $\mu\text{M}$  L-[<sup>3</sup>H]-tyrosine at 49 Ci/mmol (Sohm et al. 2003). Apparent kinetic parameters were determined from Lineweaver–Burk plots in the presence of 75 or 1.5  $\mu\text{M}$  TyrRS with transcript concentrations varying from 0.2 to 1.6  $\mu\text{M}$ . tRNAs were renatured by heating at 60°C for 90 sec in water and slow cooling to room temperature before use. Experimental errors on  $k_{\text{cat}}$  and  $K_M$  varied at most by 20%. Experiments represent an average of at least two independent experiments.

<sup>a</sup>L values correspond to losses of efficiency relative to wild-type human mt-tRNA<sup>Tyr</sup> (the value <1 corresponds to a gain in efficiency).

<sup>b</sup>Data taken from Bonnefond et al. (2005).

man mt-TyrRS of tRNAs with important sequence variations in their acceptor stems, namely, human cyt-tRNA<sup>Tyr</sup> and yeast tRNA<sup>Tyr</sup> (Fig. 2D,F; Table 1). In a next step we searched whether discriminator residue 73 is a tyrosine identity element. As found in other tyrosylation systems (Fechter et al. 2001), mutating A73 into G73 abolishes tyrosylation of human mt-tRNA<sup>Tyr</sup> ( $K_M/k_{\text{cat}}$  reduced 4800-fold; see Table 1). Summarizing, it appears that only residue A73 in the tRNA acceptor branch plays an identity role for tyrosylation of human mt-tRNA<sup>Tyr</sup>. This implies that other elements in the tRNA molecule, determinants in the acceptor branch or antideterminants, have acquired a more preponderant role in specifying tyrosylation in human mitochondria.

In an evolutionary perspective it would be surprising that the unprecedented properties of a TyrRS to charge indifferently eubacterial and eukaryal tRNA<sup>Tyr</sup> species are restricted to the human mitochondrial enzyme. Examination of mitochondrial TyrRS sequences revealed strong homologies between human and other vertebrates mt-TyrRSs. Nine putative mt-TyrRS sequences could be retrieved from the genomes of *Canis familiaris*, *Danio rerio*, *Gallus gallus*, *Mus musculus*, *Pan troglodytes*, *Ratus norvegicus*, *Sus scrofa*, *Takifugu rubripes*, and *Tetraodon nigroviridis*. The high sequence conservation in the clusters (Fig. 1) thus suggests the same dual functionality of the concerned TyrRSs toward the N1-N72 base pair. Whether the corresponding TyrRSs actually exhibit this property remains to be investigated.

## ACKNOWLEDGMENTS

We are indebted to A. Théobald-Dietrich and A. Fender for fruitful advice, and we thank M. Sissler and C. Florentz for stimulating

discussions. This work was supported by the Centre National de la Recherche Scientifique and Université Louis Pasteur (Strasbourg).

Received November 22, 2004; accepted January 14, 2005.

## REFERENCES

- Akins, R.A. and Lambowitz, A.M. 1987. A protein required for splicing group I introns in *Neurospora* mitochondria is mitochondrial tyrosyl-tRNA synthetase or a derivative thereof. *Cell* **50**: 331–345.
- Bedouelle, H. 2004. Tyrosyl-tRNA synthetases. In *Aminoacyl-tRNA synthetases* (eds. M. Ibba et al.), Chapter 12, pp. 111–124. Landes Biosciences, Georgetown, TX.
- Bedouelle, H. and Winter, G. 1986. A model of synthetase/transfer RNA interaction as deduced by protein engineering. *Nature* **320**: 371–373.
- Bedouelle, H., Guez-Ivanier, V., and Nageotte, R. 1993. Discrimination between transfer-RNAs by tyrosyl-tRNA synthetase. *Biochimie* **75**: 1099–1108.
- Beuning, P.J. and Musier-Forsyth, K. 1999. Transfer RNA recognition by aminoacyl-tRNA synthetases. *Biopolymers* **52**: 1–28.
- Bonnefond, L., Fender, A., Rudinger-Thirion, J., Giegé, R., Florentz, C., and Sissler, M. 2005. Towards the full set of human mitochondrial aminoacyl-tRNA synthetases: Characterization of AspRS and TyrRS. *Biochemistry* (in press).
- Brick, P., Bhat, T.N., and Blow, D.M. 1989. Structure of tyrosyl-tRNA synthetase refined at 2.3 Å resolution. Interaction of the enzyme with tyrosyl adenylate intermediate. *J. Mol. Biol.* **208**: 83–98.
- Celis, J.E., Hooper, M.L., and Smith, J.D. 1973. Amino acid acceptor stem of *E. coli* suppressor tRNA<sup>Tyr</sup> is a site of synthetase recognition. *Nat. New Biol.* **244**: 261–264.
- Fechter, P., Rudinger, J., Giegé, R., and Théobald-Dietrich, A. 1998. Ribozyme processed tRNA transcripts with unfriendly internal promoter for T7 RNA polymerase: Production and activity. *FEBS Lett.* **436**: 99–103.
- Fechter, P., Rudinger-Thirion, J., Théobald-Dietrich, A., and Giegé, R. 2000. Identity of tRNA for yeast tyrosyl-tRNA synthetase: Tyrosylation is more sensitive to identity nucleotides than to structural features. *Biochemistry* **39**: 1725–1733.

- Fechter, P., Rudinger-Thirion, J., Tukalo, M., and Giegé, R. 2001. Major tyrosine identity determinants in *Methanococcus jannaschii* and *Saccharomyces cerevisiae* tRNA<sup>Tyr</sup> are conserved but expressed differently. *Eur. J. Biochem.* **268**: 761–767.
- Giegé, R. 2003. Genetic code expansion. *Nat. Struct. Biol.* **10**: 414–416.
- Giegé, R., Sissler, M., and Florentz, C. 1998. Universal rules and idiosyncratic features in tRNA identity. *Nucleic Acids Res.* **26**: 5017–5035.
- Gray, M.W., Burger, G., and Lang, B.F. 2001. The origin and early evolution of mitochondria. *Genome Biol.* **2**: 1018.1–1018.5.
- Kamper, U., Kück, U., Cherniak, A.D., and Lambowitz, A.M. 1992. The mitochondrial tyrosyl-tRNA synthetase of *Podospora anserina* is a bifunctional enzyme active in protein synthesis and RNA splicing. *Mol. Cell. Biol.* **12**: 499–511.
- Kleeman, T.A., Wei, D., Simpson, K.L., and First, E.A. 1997. Human tyrosyl-tRNA synthetase shares amino acid sequence homology with a putative cytokine. *J. Biol. Chem.* **272**: 14420–14425.
- Kobayashi, T., Nureki, O., Ishitani, R., Yaremchuk, A., Tukalo, M., Cusack, S., Sakamoto, K. and Yokoyama, S. 2003. Structural basis for orthogonal tRNA specificities of tyrosyl-tRNA synthetases for genetic code expansion. *Nat. Struct. Biol.* **10**: 425–432.
- Lee, C.P. and RajBhandary, U.L. 1991. Mutants of *Escherichia coli* initiator tRNA that suppress amber codons in *Saccharomyces cerevisiae* and are aminoacylated with tyrosine by yeast extracts. *Proc. Natl. Acad. Sci.* **88**: 11378–11382.
- Nair, S., Ribas de Pouplana, L., Houman, F., Avruch, A., Shen, X., and Schimmel, P. 1997. Species-specific tRNA recognition in relation to tRNA synthetase contact residues. *J. Mol. Biol.* **269**: 1–9.
- Poirot, O., O'Toole, E., and Notredame, C. 2003. Tcoffee@igs: A web server for computing, evaluating and combining multiple sequence alignments. *Nucleic Acids Res.* **31**: 3503–3506.
- Poirot, O., Suhre, K., Abergel, C., O'Toole, E. and Notredame, C. 2004. 3DCoffee@igs: A web server for combining sequences and structures into multiple sequence alignment. *Nucleic Acids Res.* **32**: W37–W40.
- Qiu, X., Janson, C.A., Smith, W.W., Green, S.M., McDevitt, P., Johanson, K., Carter, P., Hibbs, M., Lewis, C., Chalker, A., et al. 2001. Crystal structure of *Staphylococcus aureus* tyrosyl-tRNA synthetase in complex with a class of potent and specific inhibitors. *Protein Sci.* **10**: 2008–2016.
- Quinn, C.L., Tao, N., and Schimmel, P. 1995. Species-specific microhelix aminoacylation by a eukaryotic pathogen tRNA synthetase dependent on a single base pair. *Biochemistry* **34**: 12489–12495.
- Sohm, B., Frugier, M., Brulé, H., Olszak, K., Przykorska, A., and Florentz, C. 2003. Towards understanding human mitochondrial leucine aminoacylation identity. *J. Mol. Biol.* **328**: 995–1010.
- Sprinzel, M. and Vassilenko, K.S. 2005. Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res.* **33**: D139–D140.
- Wakasugi, K., Quinn, C.L., Tao, N., and Schimmel, P. 1998. Genetic code in evolution: Switching species-specific aminoacylation with a peptide transplant. *EMBO J.* **17**: 297–305.
- Yang, X.L., Skene, R.J., McRee, D.E., and Schimmel, P. 2002. Crystal structure of a human aminoacyl-tRNA synthetase cytokine. *Proc. Natl. Acad. Sci.* **99**: 15369–15374.
- Yaremchuk, A., Kriklivyi, I., Tukalo, M., and Cusack, S. 2002. Class I tyrosyl-tRNA synthetase has a class II mode of cognate tRNA recognition. *EMBO J.* **21**: 3829–3840.

# Le système tyrosine fait entorse aux règles d' aminoacylation



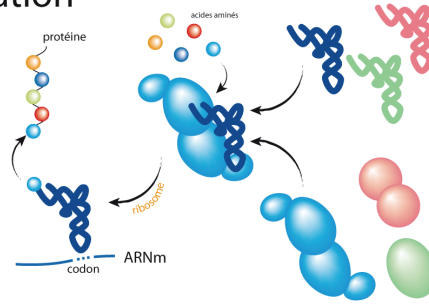
Luc BONNEFOND, Richard GIEGE et Joëlle RUDINGER-THIRION



UPR 9002 du CNRS - ARN, Institut de Biologie Moléculaire et Cellulaire, 15 rue René Descartes, F-67084 STRASBOURG cedex, email : l.bonnefond@ibmc.u-strasbg.fr

## La réaction d' aminoacylation

La **synthèse protéique** est l'étape de l'expression du génome qui permet la **traduction** de l'information génétique portée par l'**ADN**, puis l'ARN messager, en protéines. Le code génétique à quatre lettres est lu par triplets (codons) au niveau du ribosome et traduit en un langage **protéique**. Les **ARN de transfert** (ARNt) sont les **adaptateurs** qui permettent d'associer à chaque codon lu l'acide aminé correspondant. La **fidélité** de la traduction repose sur deux points : le déchiffrement correct du code génétique lors de l'association codon/anticodon, précédé de l'**aminoacylation spécifique** de l'ARNt par l'**aminoacyl-ARNt synthétase** (aaRS) correspondante. La spécificité de cette réaction repose sur la présence d'**éléments d'identité** au niveau de l'ARNt et de l'**aaRS**.



Les **ARNt** sont de petits acides nucléiques de 73 à 95 nucléotides qui se **replient en L** grâce à un réseau d'interactions tertiaires. Dans la cellule coexistent une soixantaine de ces ARNt regroupés en familles isoacceptrices d'après l'acide aminé fixé à l'extrémité 3'.

L'**aminoacylation** consiste en la **charge** d'un **acide aminé** sur l'**ARNt** homologue. Cette réaction est catalysée par l'**aaRS** correspondante.

Les **aaRSs** sont des enzymes **modulaires**, hétérogènes en séquence, organisation et oligométrie. Elles possèdent néanmoins toutes un **domaine catalytique** qui interagit avec l'extrémité acceptrice de l'ARNt.

## Un système tyrosine atypique



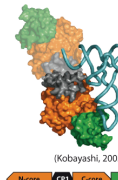
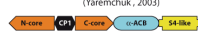
Les ARNt<sup>Tyr</sup> **eubactériens** présentent une première paire de bases **G<sub>1</sub>-C<sub>72</sub>**, qui intervient dans l'identité. Les **éléments d'identité** sont, par ordre d'importance, la base **A<sub>73</sub>**, les positions **U<sub>35</sub>** et **G<sub>34</sub>**, la **grande région variable** et la paire **G<sub>1</sub>-C<sub>72</sub>** (Himeno, 1990).



Les ARNt<sup>Tyr</sup> d'**archae** et d'**eucaryotes** présentent une première paire de bases **C<sub>1</sub>-G<sub>72</sub>**, atypique à cette position, qui joue un rôle **prépondérant** dans l'identité. Les **éléments d'identité** sont la paire **C<sub>1</sub>-G<sub>72</sub>**, la base **A<sub>73</sub>**, et dans l'anticodon **G<sub>34</sub>**, **U<sub>35</sub>** et **A<sub>36</sub>** (Fechter, 2001).



Les TyrRSs **eubactériennes** sont **homodimériques** et fixent un seul **ARNt à cheval** sur les deux sous-unités. Le domaine **catalytique** (orange) interagit avec les bases **1-72** et **73**, le domaine de **liaison à l'anticodon** (bleu) avec l'anticodon (**G<sub>34</sub>U<sub>35</sub>C<sub>36</sub>**) - et le domaine **C-terminal** (jaune) avec la **région variable**.



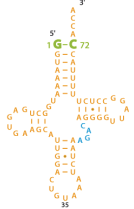
Les TyrRSs d'**archae** et d'**eucaryotes** sont également **homodimériques**. Le mode de reconnaissance globale de l'ARNt est similaire à celui des TyrRS eubactériennes, mais le **mécanisme fin** d'interaction avec l'extrémité acceptrice de l'ARNt est différent. Ainsi, toute **charge croisée** est **inefficace**.



## Les partenaires de la tyrosylation dans la mitochondrie humaine

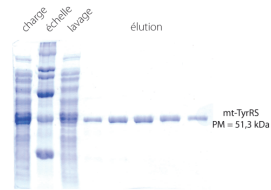
### ARNt<sup>Tyr</sup>

L'**ARNt<sup>Tyr</sup> mitochondrial humain**, riche en bases **A** et **U**, est de type **eubactérien** avec une première paire de bases **G<sub>1</sub>-C<sub>72</sub>**. Cependant cet ARNt présente une **petite région variable** tout comme les ARNt<sup>Tyr</sup> **eucaryotiques**. L'ARNt<sup>Tyr</sup> a été synthétisé par **transcription in vitro** à l'aide de la T7 ARN polymérase.



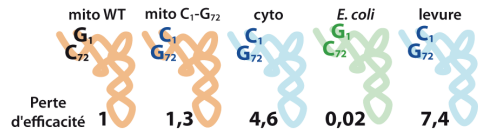
### TyrRS

La **TyrRS mitochondriale humaine** a été préalablement **identifiée** dans les banques de données génomiques, puis sa séquence amplifiée par **pCR** a été clonée dans un vecteur d'expression permettant l'ajout d'un **tag His6** en C-terminal. La protéine surexprimée dans *E. coli* a été purifiée par **chromatographie d'affinité**. (Bonnefond, Biochemistry, 2005)



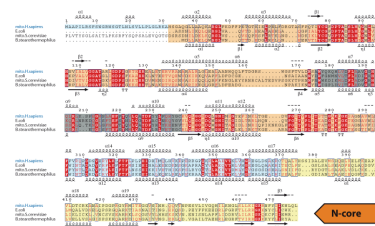
## Le rôle de la paire 1-72

Le **système** de tyrosylation mitochondriale humaine **reconstitué in vitro** est **actif**, mais moins efficace qu'un système eubactérien classique. L'influence de la paire de bases 1-72 a été testée dans différents ARNt<sup>Tyr</sup>.



Les **paramètres cinétiques** de l' aminoacylation de différents ARNt<sup>Tyr</sup> transcripts *in vitro* ou natifs ont été mesurés. Les **pertes d'efficacité** sont **faibles** pour tous les ARNt et indiquent une **absence de spécificité** de la TyrRS mitochondriale humaine vis-à-vis de la **première paire de bases 1-72** (Bonnefond, RNA, 2005).

### Alignement de TyrRS eubactériennes avec la mt-TyrRS humaine.



La TyrRS mitochondriale humaine a une **organisation modulaire** semblable à celle des TyrRS **eubactériennes**. Cependant l'**analyse détaillée** de sa séquence au niveau de la zone d'interaction avec la paire de bases 1-72 ne permet pas de prévoir le mécanisme d'interaction de celle-ci avec son ARNt homologue.



## Conclusions et perspectives

La TyrRS mitochondriale humaine est la première TyrRS connue sans spécificité pour la paire de bases 1-72 de l'ARNt<sup>Tyr</sup>. Elle semble capable d' aminoacyler les ARNt<sup>Tyr</sup> indépendamment de la nature de cette paire de bases.

Le mécanisme de reconnaissance de l'ARNt<sup>Tyr</sup> mitochondrial se différencie de celui décrit pour les autres systèmes de tyrosylation.

Les différences fonctionnelles entre phyla permettraient la mise au point d'inhibiteurs bloquant spécifiquement la réaction de tyrosylation eubactérienne.

### Bibliographie :

- Himeno et coll. (1990) Nucleic Acid Res. **18**, 6815-6819
- Fechter et coll. (2001) Eur. J. Biochem. **268**, 761-767
- Yaremchuk et coll. (2002) EMBO J. **21**, 3829-3840
- Kobayashi et coll. (2003) Nat. Struct. Biol. **10**, 425-432
- Bonnefond et coll. (2005) Biochemistry, **44**, 4805-4816
- Bonnefond et coll. (2005) RNA, **11**, 558-562

Figure 14. Affiche pour le "Congrès annuel de la SFBBM" et le "32<sup>ème</sup> Forum des Jeunes Chercheurs" (Nantes, 24-26 Octobre 2005)

# Mitochondrial TyrRS charges eukaryotic tRNA<sup>Tyr</sup>



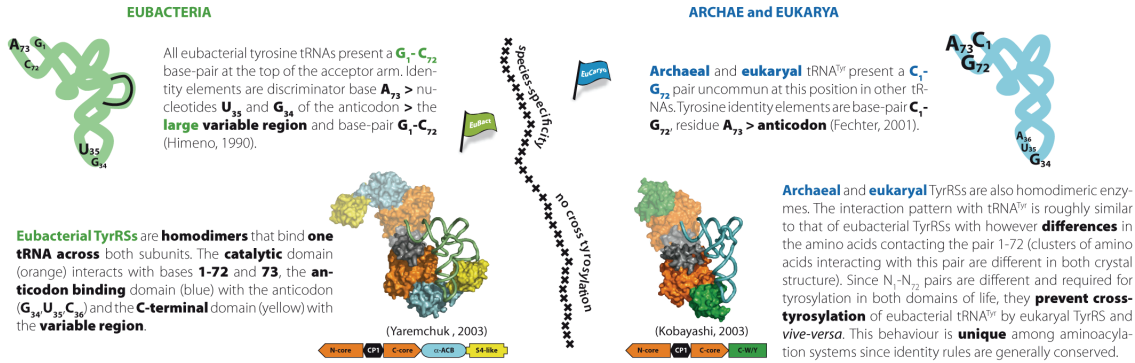
Luc BONNEFOND, Magali FRUGIER, Richard GIEGE and Joëlle RUDINGER-THIRION



UPR 9002 du CNRS - ARN, Institut de Biologie Moléculaire et Cellulaire, 15 rue René Descartes, F-67084 STRASBOURG cedex, email : l.bonnefond@ibmc.u-strasbg.fr



## What is known about tyrosine identity ?



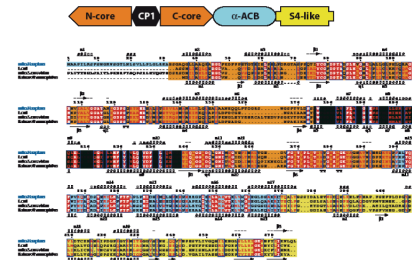
## ...and in human mitochondria ?

### Synthesis of human mt-tRNA<sup>Tyr</sup> and TyrRS

**Human mitochondrial tRNA<sup>Tyr</sup>** presents numerous **A** and **U** base-pairs (as usual in mt-tRNAs). It is of **eubacterial** type with a base-pair **G<sub>1</sub>-C<sub>72</sub>**. However, this tRNA shares a **small variable region**, as all **eukaryotic** tRNA<sup>Tyr</sup>. Noticeable is the absence of several conserved residues in the T- and D-loops responsible of the long range interactions as described for many mt-tRNAs. This tRNA (and variants) were produced by *in vitro* transcription using T7 RNA polymerase.



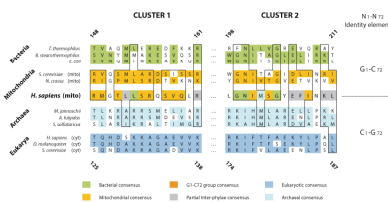
**Human mt-TyrRS sequence** was searched in genomic databanks, **identified** and amplified by **PCR Cloning** was performed in an expression vector (pQE70) and the corresponding protein was purified (Bonnefond, 2005a). Human mt-TyrRS (i) shares 458 aa (PM 51,3 kDa) (ii) is a **class I** synthetase with HWGH and KLGKS consensus sequences, (iii) shares a **modular organization** similar to that of other TyrRSs, (iv) presents an anticodon binding domain **comparable** to that of **eubacterial** TyrRSs ( $\alpha$ -helical) that distinguishes from archaeal and eukaryotic TyrRSs.



### Specificity of mt-TyrRS towards N<sub>1</sub>-N<sub>72</sub> base-pair

#### Sequence analysis

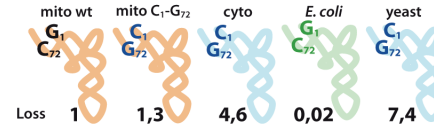
Previous comparisons of TyrRS sequences revealed a **correlation** between two sequence clusters of the catalytic domain and the N<sub>1</sub>-N<sub>72</sub> identity (Nair, 1997). Analysis of human mt-TyrRS sequence of clusters 1 & 2 emphasizes **new sequence characteristics** and a **mosaic character** not found in earlier studied TyrRSs. Thus, it was not able to predict the recognition mode with 1-72 pair. It implies a **different recognition pattern** of the tRNA<sup>Tyr</sup> at the level of G<sub>1</sub>-C<sub>72</sub> base-pair.



#### Generalization to other mt-TyrRSs ?

These functional and sequential properties **might be generalized** to other mitochondrial TyrRSs from **higher eukarya** since they share strong sequence similarities with human mt-TyrRS. Nevertheless, sequence similarities with eubacterial TyrRSs suggest a general interaction scheme close to that found in the *Thermus thermophilus* system.

#### Functional analysis



Kinetic parameters for tyrosylation of different tRNA<sup>Tyr</sup> molecules were determined *in vitro*. Catalytic activity obtained with the homologous transcript is less efficient than that obtained with native *E. coli* molecules (already described for many mt systems). Surprisingly, variants of the first base-pair exhibit **no strong losses** of specificity as compared to the wt. This demonstrates the **non-involvement** of this pair in mitochondrial tyrosine identity and provides the **first example of a TyrRS lacking specificity towards N1-N72** (Bonnefond, 2005b).

#### Aminoacylation conditions

- 50 mM HEPES-NaOH pH 7.6
- 25 mM KCl
- 12 mM MgCl<sub>2</sub>
- 2.5 mM ATP
- 0.2 mg/ml BSA
- 1 mM spermidine
- 10  $\mu$ M [H]-Tyr
- 37°C

#### References

- Bonnefond *et al.* (2005a) *Biochemistry*, **44**, 4805-4816.
- Bonnefond *et al.* (2005b) *RNA*, **11**, 558-562.
- Fechter *et al.* (2001) *Eur. J. Biochem.*, **268**, 761-767.
- Himeno *et al.* (1990) *Nucleic Acid Res.*, **18**, 6815-6819.
- Kobayashi *et al.* (2003) *Nat. Struct. Biol.*, **10**, 425-432.
- Nair *et al.* (1997) *J. Mol. Biol.*, **269**, 1-9.
- Yaremchuk *et al.* (2002) *EMBO J.*, **21**, 3829-3840.

Figure 15. Affiche pour le "21<sup>st</sup> International tRNA Workshop", (Bangalore, Inde, 2-7 Décembre 2005)

## 2. Influence de la nature de la première paire de bases N1-N72

La capacité de la TyrRS mitochondriale à charger des ARNt<sup>Tyr</sup> présentant une première paire de bases G1-C72 ou C1-G72 a mis en évidence son absence de discrimination entre des ARNt<sup>Tyr</sup> de type eubactériens ou archaea/eucaryotes. Pour confirmer la perte totale de spécificité de la mt-TyrRS vis-à-vis de la première paire de base des ARNt<sup>Tyr</sup>, l'aminoacylation de l'ARNt<sup>Tyr</sup> mitochondrial muté en A1-U72 et U1-A72 a été testée. Il est intéressant de noter que la paire de bases A1-U72 est retrouvée exclusivement chez des ARNt<sup>Tyr</sup> mitochondriaux (10 d'animaux, celui d'*Arabidopsis thaliana* et 4 d'organismes unicellulaires) et l'ARNt<sup>Tyr</sup> apicoplastique de *Plasmodium falciparum* (Sprinzl et Vassilenko, 2005). Quant à la paire U1-A72, nous pensons que sa présence chez un seul organisme pourrait être le fait d'une erreur de séquençage ou d'assignation plutôt qu'une réalité biologique.

Transcrit mt-ARNt <sup>Tyr</sup>	type	$K_M$ ( $\mu\text{M}$ )	$k_{\text{cat}}$ ( $10^{-3} \cdot \text{s}^{-1}$ )	$k_{\text{cat}}/K_M$ ( $10^{-3} \cdot \text{s}^{-1} \cdot \mu\text{M}^{-1}$ )	perte
G1-C72	bactérien/mito	4,8	46	9,6	1
C1-G72	archaea/euk	0,48	3,4	7,1	1,3
A1-U72	mito	4,5	4,2	0,9	10
U1-A72		2,1	4,2	2	5

**Tableau 3. Paramètres cinétiques de l'aminoacylation de transcrits de l'ARNt<sup>Tyr</sup> mitochondrial muté en N1-N72**

Les différents mutants de la première paire de bases de l'ARNt<sup>Tyr</sup> mitochondrial sont aminoacylés par la mt-TyrRS avec une efficacité catalytique similaire, ce qui confirme l'absence totale de spécificité vis-à-vis de la nature de cette paire.

Le modèle d'interaction ARNt<sup>Tyr</sup>/TyrRS préalablement décrit par Bedouelle pour *B. stearothermophilus* (Bedouelle et Winter, 1986), qui est l'organisme dont la TyrRS est la plus proche en séquence et en structure de la mt-TyrRS, soulignait l'importance de la structuration correcte avec des paires de bases canoniques de l'extrémité de la branche acceptrice. Il serait intéressant de vérifier si le système mitochondrial requiert aussi cette structuration de l'ARNt<sup>Tyr</sup> pour l'activité d'aminoacylation ou si sa flexibilité "mitochondriale" lui permet d'adapter de plus grandes variations de séquences avec notamment une première paire de bases N1-N72 non appariée.



### **3. Rôle accru de la base discriminatrice et de l'anticodon**

De manière générale, les éléments d'identité des ARN<sup>Tyr</sup> sont constitués de la première paire de bases N1-N72, la base discriminatrice A73, les trois positions de l'anticodon GUA et la grande région variable pour les ARN<sup>Tyr</sup> eubactériens. L'absence d'implication de la première paire de bases dans le cas de la mitochondrie humaine a été testée et confirmée. Pour déterminer le jeu d'identité complet de l'ARN<sup>Tyr</sup> mitochondrial humain, les autres positions citées ont aussi été mutées, à l'exception de la région variable qui est de taille réduite pour l'ARN<sup>Tyr</sup> mitochondrial humain. Les résultats des tests fonctionnels réalisés sur les molécules mutées ont été présentés lors de la "5<sup>ème</sup> rencontre sifrARN" (Arcachon, Octobre 2004) puis publiés dans un article de revue sur les systèmes d'aminacylation spécifiques de la tyrosine paru en 2005 dans un numéro spécial de la revue "Biochimie" consacré à ce congrès.

Article de revue n°1

#### **Evolution des systèmes d'aminacylation ARN<sup>tTyr</sup>/TyrRS**

Luc Bonnefond, Richard Giegé et Joëlle Rudinger-Thirion

**2005**

Biochimie, 87 : 873-883

Cet article est présenté en pages 49 à 59 de l'**Introduction**.

Dans un souci d'homogénéité, les paramètres exacts de l'aminacylation des molécules testées n'ont pas été précisés dans l'article, mais sont rapportés ici.

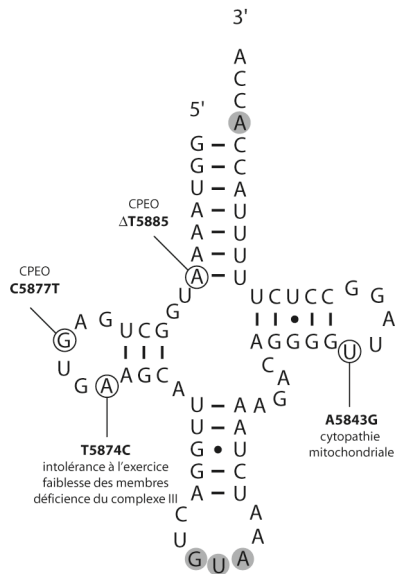
<b>transcrits</b>	<b>localisation</b>	$K_M$	$k_{cat}$	$k_{cat}/K_M$	<b>perte</b>
<b>mt-ARNt<sup>Tyr</sup></b>		( $\mu\text{M}$ )	( $10^{-3} \cdot \text{s}^{-1}$ )	( $10^{-3} \cdot \text{s}^{-1} \cdot \mu\text{M}^{-1}$ )	
sauvage		4,8	46	9,6	1
G34C	anticodon	44	0,1	0,002	5000
U35G	anticodon	1,9	0,01	0,005	2000
A36C	anticodon	2	0,01	0,005	2000
A73G	base discriminatrice	24	0,05	0,002	5000

**Tableau 4. Paramètres cinétiques de l'aminacylation de transcrits de l'ARNt<sup>Tyr</sup> mitochondrial mutés au niveau du triplet anticodon et de la base discriminatrice**

Chacune des positions testées joue un rôle très fort dans le système tyrosine mitochondrial humain. Dans tous les cas, l'efficacité catalytique de la réaction est sévèrement affectée. Pour les autres systèmes de tyrosylation décrits seuls un ou deux nucléotides de l'anticodon jouent un rôle fort dans l'identité de l'ARNt correspondant. Ici, les trois positions jouent un rôle prépondérant. Ce phénomène pourrait compenser la perte de spécificité vis-à-vis de la première paire de bases, mais aussi suite à la réduction des contacts entre le domaine C-terminal S4-like et l'ARNt<sup>Tyr</sup> du fait de la petite taille de sa région variable.

#### **4. Mutations pathogéniques de l'ARNt<sup>Tyr</sup> mitochondrial**

Le jeu d'éléments d'identité est-il composé seulement de quatre nucléotides situés aux extrémités de la structure en "L" de l'ARNt<sup>Tyr</sup> ? Pour y répondre nous avons procédé de manière rationnelle en testant tout d'abord les positions de l'ARNt<sup>Tyr</sup> mitochondrial impliquées dans des pathologies et donc potentiellement importantes pour la réaction d'aminacylation.



**Figure 16. Mutations pathologiques de l'ARNt<sup>Tyr</sup> mitochondrial humain**

Les positions correspondant à des mutations pathogéniques (numérotation selon le gène) sont entourées et la pathologie correspondante indiquée. La position des éléments d'identité est rappelée par un ombrage gris. CPEO, Chronical Progressive External Ophthalmoplegia.

D'après l'analyse de la séquence du génome mitochondrial de patients, quatre mutations dans le gène de l'ARNt<sup>Tyr</sup> mitochondrial humain ont été corrélées à des pathologies de type myopathie ou cytopathie. Il s'agit d'une délétion d'un nucléotide dans le bras accepteur ( $\Delta A7$  - Raffelsberger et al., 2001) et de mutations des résidus 15 (A15G - Sahashi et al., 2001) et 22 dans le bras D (A22G - Pulkes et al., 2000) et du résidu 55 dans la boucle T (T55C - Scaglia et al., 2003). Ces 3 mutations ont été introduites dans des transcrits de l'ARNt<sup>Tyr</sup> mitochondrial pour reproduire les séquences observées chez les patients. Les paramètres cinétiques de leur aminoacylation par la TyrRS mitochondriale ont été déterminés. Parallèlement, des expériences préliminaires de détermination de structure par l'utilisation de sondes enzymatiques ont été réalisées pour détecter d'éventuelles modifications de structure de l'ARNt induites par ces mutations.

**4.1. Effet des mutations sur l'aminacylation par la TyrRS mitochondriale**

Les effets de ces mutations sur la réaction d'aminacylation diffèrent suivant les positions mutées (**tableau 7**). Alors que la mutation en position 55 n'entraîne pas de perte d'activité, celle en position 15 provoque une perte d'activité de 40 fois et celle en position 22 une forte perte d'activité de 600 fois. L'absence d'effet d'une mutation sur la réaction d'aminacylation ne remet pas en cause son implication dans une pathologie. La mutation peut perturber d'autres mécanismes impliquant l'ARNt (synthèse, maturation, modification,



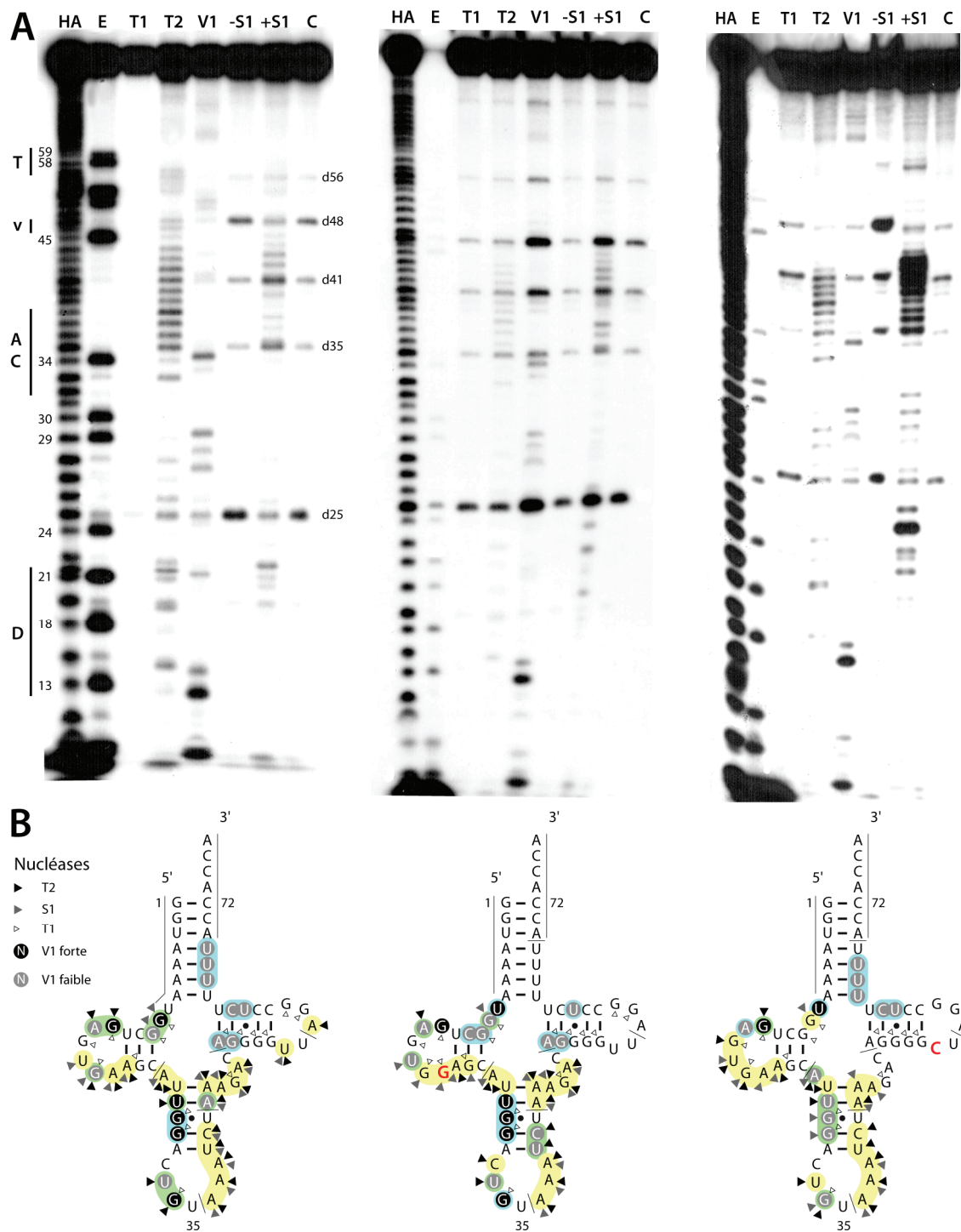
transport...). Ici, comme pour les quatre éléments d'identité identifiés, l'effet des mutations se répercute principalement sur la vitesse de la réaction ce qui semble exclure un effet direct des mutations sur la structure de l'ARNt<sup>Tyr</sup> mitochondrial. Ceci peut s'expliquer par la localisation des mutations situées exclusivement dans des boucles. Néanmoins, de manière générale, la structuration des ARNt mitochondriaux est mise en cause pour expliquer la perte d'efficacité au niveau de la réaction d'aminacylation (Levinger et al., 2004).

<b>transcrits</b> <b>mt-ARNt<sup>Tyr</sup></b>	<b>localisation</b>	<b><math>K_M</math></b> ( $\mu\text{M}$ )	<b><math>k_{\text{cat}}</math></b> ( $10^{-3} \cdot \text{s}^{-1}$ )	<b><math>k_{\text{cat}}/K_M</math></b> ( $10^{-3} \cdot \text{s}^{-1} \cdot \mu\text{M}^{-1}$ )	<b>perte</b>
sauvage		4,8	46	9,6	1
A15G	boucle D	0,8	0,2	0,25	40
A22G	boucle D	1,6	0,03	0,016	600
T55C	boucle T	2,2	12	5,4	1,8

**Tableau 5. Paramètres cinétiques de l'aminacylation de transcrits de l'ARNt<sup>Tyr</sup> mitochondrial mutés à des positions pathogéniques**

#### **4.2. Détermination de la structure des mutants pathogéniques du mt-ARNt<sup>Tyr</sup> mitochondrial humain**

Pour pouvoir confirmer l'absence de perturbation de la structure des transcrits de l'ARNt<sup>Tyr</sup> mitochondrial suite aux mutations introduites, la structure en solution des molécules sauvage et mutées aux positions 22 et 55 a été déterminée (**figure 17**). Nous avons choisi une mutation sans effet apparent sur l'aminacylation (55) et une autre avec un effet significatif (22), ceci dans l'idée de corréliser ces différences d'activités à des perturbations structurales. Les structures en solution ont été réalisées avec 4 sondes enzymatiques : la nucléase T1 qui hydrolyse l'ARN au niveau des résidus G, les nucléases T2 et S1 qui sont spécifiques des régions simples brins et la nucléase V1 hydrolyse les régions double-brins.

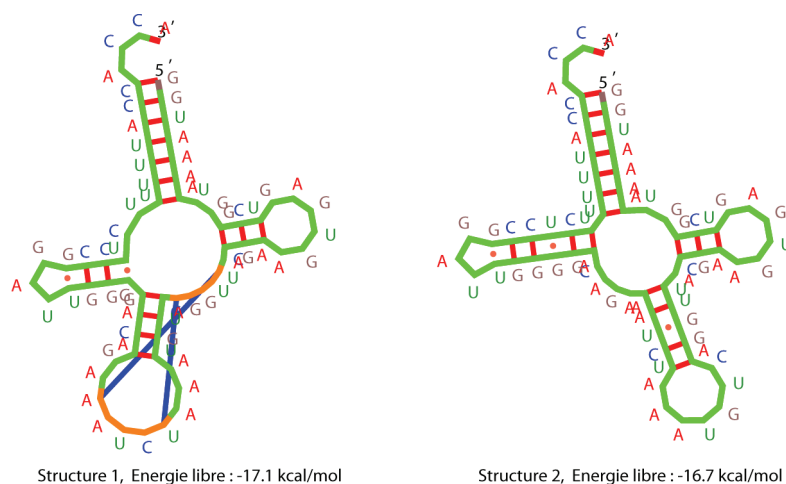


**Figure 17. Cartographie en solution de transcrits de l'ARNt<sup>Tyr</sup> mitochondrial humain**

(A) Autoradiogrammes de cartographie de transcrits de l'ARNt<sup>Tyr</sup> mitochondrial humain sauvage, et mutés (A22G et T55C) sur gels dénaturants 12%. **HA**, échelle d'hydrolyse alcaline ; **E**, échelle d'hydrolyse à la RNase T1 en conditions dénaturantes ; **C**, contrôle en absence de sonde ; **-S1**, effet du ZnCl<sub>2</sub> présent dans le tampon de la nucléase S1. Les autres colonnes correspondent aux hydrolyses par les différentes nucléases. Les numéros des G et des positions de dégradation spontanée (**d**) sont indiqués, de même que les boucles **T** (**T**), anticodon (**AC**), **D** (**D**) et la région variable (**v**).

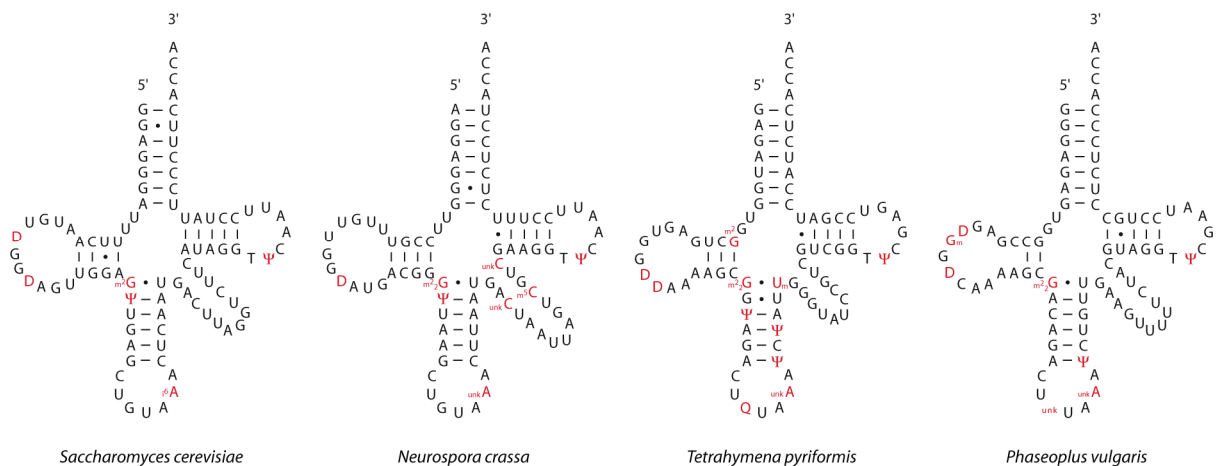
(B) Résultats de cartographie de la structure en feuille de trèfle de l'ARNt<sup>Tyr</sup>. Les lignes aux extrémités 5' et 3' indiquent les régions non analysées. Les surfaces jaunes correspondent aux nucléotides coupés par les nucléases spécifiques des régions simple brin, les surfaces bleues pour les nucléotides coupés par les nucléases spécifiques des régions double-brins, et les nucléotides coupés par les deux types d'enzyme sont soulignés en vert. Les traits indiquent les dégradations spontanées.

Parallèlement aux ARNt<sup>Tyr</sup> mutés, la structure de l'ARNt<sup>Tyr</sup> mitochondrial sauvage a aussi été déterminée afin de servir de référence, mais aussi pour connaître la structure intrinsèque de cet ARNt mitochondrial (**figure 17**). Le profil de digestion à la nucléase V1 nous révèle que les branches acceptrice, anticodon et T sont structurées en double brin. Les profils obtenus avec les nucléases S1 et T1 confirment la présence de zones simples brins au niveau des boucles D, anticodon et T. On observe aussi un nombre important de nucléotides sensibles aux sondes ciblant les régions simple et double brins. Ceci suggère une structure tertiaire relâchée, déjà suspectée par l'absence de la plupart des résidus conservés impliqués dans le réseau d'interactions tertiaires classique des ARNt. Cet ARNt pourrait aussi présenter des conformations alternatives comme le suggère la prédiction réalisée avec Kinefold (Xayaphoummine et al., 2005).



**Figure 18. Prédiction de structures secondaires de l'ARNt<sup>Tyr</sup> mitochondrial humain**

Ce logiciel prédit une conformation alternative (**figure 18**, structure 1) légèrement plus stable que la conformation classique (structure 2) et qui fait intervenir des appariements de type Watson-Crick entre les nucléotides 32-35 de la boucle anticodon et 47-50 de la région variable. Cette prédiction n'est pas en accord avec les profils de digestions aux différentes sondes et ne serait donc qu'un artefact. Rappelons que les molécules testées sont dépourvues de modifications post-transcriptionnelles qui contribuent probablement à la structuration et à la stabilité des molécules. La position de ces modifications n'est pas connue pour l'ARNt<sup>Tyr</sup> mitochondrial humain, mais a été déterminée pour les levures *S. cerevisiae* (Sibler et al., 1983) et *N. crassa* (Heckman et al., 1979), le protozoaire *Tetrahymena pyriformis* (Schnare et al., 1985) et le haricot commun *Phaseolus vulgaris* (Maréchal et al., 1985).



### Figure 19. Séquences et structures secondaires d'ARNt<sup>Tyr</sup> mitochondriaux

Les modifications post-transcriptionnelles sont indiquées en rouge. Séquences extraites de la tRNA database (Sprinzl et Vassilenko, 2005). D, dihydrouridine ; m<sup>2</sup>G, N<sup>2</sup>-methylguanosine, Gm, 2'-O-methylguanosine ; Ψ, pseudouridine ; i<sup>6</sup>A, N<sup>6</sup>-isopennyladienosine ; m<sup>2,2</sup>G, N<sup>2</sup>,N<sup>2</sup>-dimethylguanosine ; unkA, adénosine modifiée non identifiée ; unkC, cytidine modifiée inconnue ; m<sup>5</sup>C, 5-méthylcytidine ; Q, queuosine ; Um, 2'-O-méthyluridine ; unk, nucléotide modifié non identifié.

On retrouve pour ces 4 ARNt<sup>Tyr</sup> des modifications post-transcriptionnelles au niveau des bras D et anticodon, ainsi que la pseudouridine conservée du bras T (**figure 19**). Malheureusement pour notre analyse, ces ARNt mitochondriaux sont très proches des ARNt classiques et conservent notamment la plupart des nucléotides conservés impliqués dans le repliement en "L" de l'ARNt. A l'opposé, le mt-ARNt<sup>Tyr</sup> humain présente des boucles D et T de taille très réduite et n'a conservé que quelques nucléotides du réseau d'interactions tertiaires, ce qui expliquerait sa flexibilité. D'après la position des modifications des 4 ARNt<sup>Tyr</sup> mitochondriaux séquencés, on peut penser que dans le cas du mt-ARNt<sup>Tyr</sup> humain les modifications permettraient de stabiliser la branche anticodon, les interactions entre boucle D et T et la reconnaissance par la mt-TyrRS de la boucle anticodon.

Dans le cas des ARNt<sup>Tyr</sup> mutés en position 22 et 55 dans les boucles D et T, respectivement, aucune modification significative de leurs profils de digestions n'apparaît (**figure 17**). Le mutant de la position 55 semble présenter une structure légèrement plus relâchée, ce qui s'expliquerait si le résidu en position 55 est impliqué dans une interaction tertiaire. Cette éventuelle relaxation n'entraîne pourtant pas de perte d'efficacité d'aminocyclation pour ce mt-ARNt<sup>Tyr</sup> alors qu'à l'opposé la mutation du résidu en position 22 entrave l'aminocyclation sans perturber la structure du mt-ARNt<sup>Tyr</sup>. Pourtant ce résidu en position a été supposé impliqué dans une interaction tertiaire ([13-22]-46 entre le bras D et la région variable; Helm et al., 2000). Le résidu A22 pourrait donc constituer un élément du

jeu d'identité de l'ARNt<sup>Tyr</sup> mitochondrial humain. Ces résultats sont néanmoins à confirmer en réalisant des expériences d'empreintes de la mt-TyrRS sur ces molécules.

## 5. Le jeu d'identité est-il complet ?

La validation d'un jeu d'identité consiste à le transplanter dans un ARNt hôte. L'acquisition d'une nouvelle spécificité par l'ARNt hôte confirme la complétude du jeu d'identité. L'étude parallèle du système mt-ARNt<sup>Asp</sup>/mt-AspRS humain par Aurélie Fender au laboratoire a été mise à profit pour la transplantation des 4 éléments d'identité du mt-ARNt<sup>Tyr</sup> dans le mt-ARNt<sup>Asp</sup> et *vice versa*. Néanmoins, ces manipulations se sont révélées inefficaces puisque aucun des ARNt hôte n'a acquis la spécificité attendue. Ce résultat indique soit que le jeu d'éléments d'identité du mt-ARNt<sup>Tyr</sup> humain est incomplet, soit que le choix de l'ARNt<sup>Asp</sup> comme hôte n'était pas approprié pour la transplantation. En effet, les deux mt-ARNt sont codés par des brins différents du génome mitochondrial et possèdent des compositions en bases différentes. De plus, la taille des boucles D et T n'est pas la même d'où la nécessité d'un affinement structural pour la présentation optimale du jeu d'identité dans l'ARNt hôte (Perret et al., 1992). Le mt-ARNt<sup>Asp</sup> pourrait aussi contenir un antidéterminant pour la mt-TyrRS ou ne pas posséder les éléments "permissifs" (Frugier et al., 1998) nécessaires à l'expression du jeu d'identité. Un des objectifs consisterait à manipuler ces identités dans des contextes structuraux sensiblement différents. Enfin, il est possible que les systèmes d'aminoacylation mitochondriaux prennent en compte des éléments des mt-ARNt qui ne sont pas considérés dans les systèmes d'aminoacylation "classiques", ce qui leur permettrait de conserver la fidélité de l'aminoacylation en dépit de la relaxation apparente de l'identité dans ces systèmes. Jusqu'à présent le seul cas décrit de transplantation positive d'éléments d'identité dans le cas de systèmes d'aminoacylation mitochondriaux concerne le système alanine de drosophile. Dans ce système, la transplantation de la seconde paire de bases G2-C71 dans une minihélice tryptophane a conféré à celle-ci une spécificité alanine (Lovato et al., 2001).

## **C. Caractérisation structurale de la TyrRS mitochondriale**

Pour essayer de mieux comprendre les particularités fonctionnelles de la TyrRS mitochondriale humaine, nous avons entrepris de caractériser cette TyrRS par une approche de biologie structurale. Nous avons choisi de déterminer sa structure par diffraction des rayons X, une technique connue et d'usage fréquent au laboratoire. Nous nous sommes néanmoins heurtés à un problème de production de protéine apte à cristalliser. En effet la TyrRS mitochondriale humaine purifiée par chromatographie d'affinité n'est pas homogène et stable. La modification des conditions de purification a permis d'améliorer la stabilité, mais pas l'homogénéité. L'introduction de diverses étapes de purification supplémentaire (résine échangeuse d'ions et gel-filtration), n'a pas permis d'isoler une protéine homogène.

### **1. Obtention de cristaux de TyrRS mitochondriale exploitables pour la résolution de la structure**

Une forme stable et homogène de la protéine a été obtenue après suppression du domaine C-terminal homologue à la protéine ribosomale S4 (S4-like) qui est probablement très mobile en absence d'ARNt par rapport aux domaines catalytique et de liaison à l'anticodon. Cette stratégie avait déjà été utilisée pour plusieurs TyrRS eubactériennes (Brick et Blow, 1987 ; Qiu et al., 2001) et la TyrRS mitochondriale de *N. crassa* (Paukstelis et al., 2005). La protéine recombinante tronquée obtenue est homogène, stable et a pu être utilisée pour faire pousser des cristaux de mt-TyrRS. L'obtention de ces cristaux et des clichés de diffractions des rayons X correspondants ont été présentés dans un article de la revue "Acta Crystallographica Section F" en 2007.

Article n°4

## **La tyrosyl-ARNt synthétase : premiers cristaux d'une aminoacyl-ARNt synthétase mitochondriale humaine**

Luc Bonnefond, Magali Frugier, Elodie Touzé, Bernard Lorber, Catherine Florentz, Richard Giegé, Joëlle Rudinger-Thirion et Claude Sauter

**2007**

Acta Crystallographica **F63**, 338–341

La tyrosyl-ARNt synthétase mitochondriale humaine ainsi qu'une version tronquée dépourvue de son domaine carboxy-terminal S4-like ont été purifiées et cristallisées. Seule la version tronquée, capable d'activer la tyrosine et de la transférer sur l'ARNt<sup>Tyr</sup> d'*Escherichia coli*, a permis l'obtention de cristaux exploitables pour la résolution de la structure. Ces cristaux tétragonaux, appartenant au groupe d'espace P4<sub>3</sub>2<sub>1</sub>2, ont été obtenus en présence de PEG4000 comme agent cristallisant. Ils diffractent les rayons X à une résolution de 2.7 Å. Des jeux de données complets ont pu être enregistrés et phasés par remplacement moléculaire.

Luc Bonnefond, Magali Frugier,  
Elodie Touzé, Bernard Lorber,  
Catherine Florentz, Richard  
Giegé,\* Joëlle Rudinger-Thirion  
and Claude Sauter

Département 'Machineries Traductionnelles',  
Architecture et Réactivité de l'ARN, Université  
Louis Pasteur de Strasbourg, CNRS, IBMC,  
15 Rue René Descartes, 67084 Strasbourg,  
France

Correspondence e-mail:  
r.giege@ibmc.u-strasbg.fr

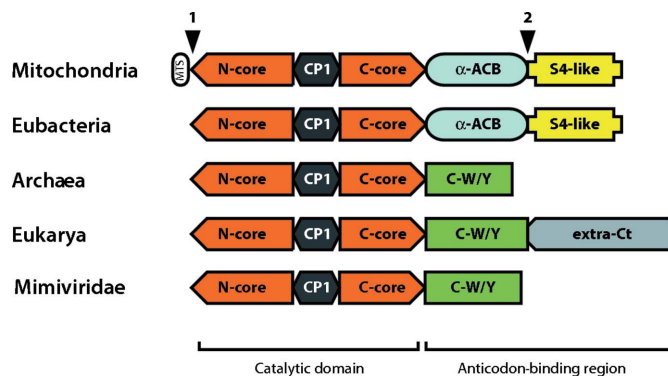
Received 31 January 2007  
Accepted 16 March 2007

## Tyrosyl-tRNA synthetase: the first crystallization of a human mitochondrial aminoacyl-tRNA synthetase

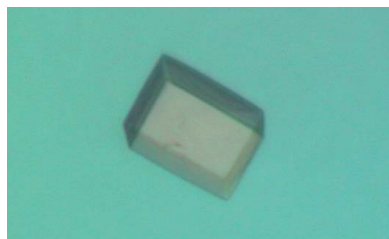
Human mitochondrial tyrosyl-tRNA synthetase and a truncated version with its C-terminal S4-like domain deleted were purified and crystallized. Only the truncated version, which is active in tyrosine activation and *Escherichia coli* tRNA<sup>Tyr</sup> charging, yielded crystals suitable for structure determination. These tetragonal crystals, belonging to space group  $P4_32_12$ , were obtained in the presence of PEG 4000 as a crystallizing agent and diffracted X-rays to 2.7 Å resolution. Complete data sets could be collected and led to structure solution by molecular replacement.

### 1. Introduction

Aminoacyl-tRNA synthetases (aaRSs) are ancient enzymes that ensure the correct attachment of amino acids on tRNA 3' ends. Each amino acid has its specific synthetase; there are therefore 20 aaRSs in nature, divided into two classes of ten based on structural and functional features of their catalytic domain (Cusack *et al.*, 1990; Eriani *et al.*, 1990). Tyrosyl-tRNA synthetases (TyrRSs) are modular proteins that belong to the first class of synthetases, but have dimeric organization and a class II mode of tRNA recognition (Yaremchuk *et al.*, 2002). TyrRS from *Bacillus stearothermophilus* was the first synthetase to have its structure solved (Bhat *et al.*, 1982) and at present TyrRSs, with structures of representatives of the three kingdoms of life (Bedouelle, 2005) and even from the mimivirus infecting the protist *Acanthamoeba polyphaga* (Abergel *et al.*, 2005), are among the best-known aaRSs (Fig. 1). However, many features related to the function, phylogeny and structure of this family of enzymes remain to be deciphered. As far as function is concerned, TyrRSs are peculiar in the way they specify cognate tRNA recognition. While major identity elements in tRNA for recognition by



**Figure 1**  
Structural organization of mitochondrial TyrRSs compared with their homologues in the three domains of life and in mimivirus. Each structural domain is displayed in a different colour. Abbreviations are MTS, mitochondrial target sequence; N-core and C-core, N and C parts of the catalytic domain, respectively; CP1, connective peptide;  $\alpha$ -ACB,  $\alpha$ -helical anticodon-binding domain; C-W/Y, C-terminal domain homologue to TrpRS; S4-like, ribosomal protein S4-like domain; extra-Ct, additional C-terminal domain. The full-length version of human mt-TyrRS (458 residues) was overexpressed without the MTS extension (arrow 1) and the truncated enzyme (356 residues) was overexpressed after further removal of the C-terminal S4-like domain (arrow 2). Notice that only one mimivirus is known; it codes for a TyrRS that resembles archaeal TyrRSs (Abergel *et al.*, 2005).





aaRSs have mostly been conserved in evolution (Giegé *et al.*, 1998), this is not the case in the tyrosine system. The major determinant, namely base pair N1–N72 at the top of the tRNA<sup>Tyr</sup>-accepting branch, is G–C in bacteria/mitochondria but C–G in archaea/eukarya. Moreover, for reasons that are still unclear, base pair G1–C72 has lost its functional role in human mitochondria and most likely also in other vertebrate mitochondria (Bonnefond, Frugier *et al.*, 2005). To date, structural knowledge of mitochondrial aaRSs is limited to only two enzymes, namely TyrRS from *Neurospora crassa* (Paukstelis *et al.*, 2005) and bovine SerRS (Chimnaronek *et al.*, 2004). Our interest in mitochondrial aaRSs (Bonnefond, Fender *et al.*, 2005) also stems from the structural and functional peculiarities of mitochondrial tRNA aminoacylation systems compared with their cytoplasmic homologues (Sissler *et al.*, 2005) and, in the particular case of human mitochondria, from the correlations between pathologies and defects in the translational machinery (Florentz *et al.*, 2003; Jacobs & Turnbull, 2005). For all these reasons, we focused our attention on TyrRS from *Homo sapiens* mitochondria, a protein that shows 37% sequence identity with the most closely related TyrRS of known structure, namely that from *B. stearothermophilus*. Here, we report the crystallization and X-ray crystallographic studies of this human TyrRS in its full-length (human mt-TyrRS) and C-terminally truncated (human mt-TyrRS- $\Delta$ S4) versions.

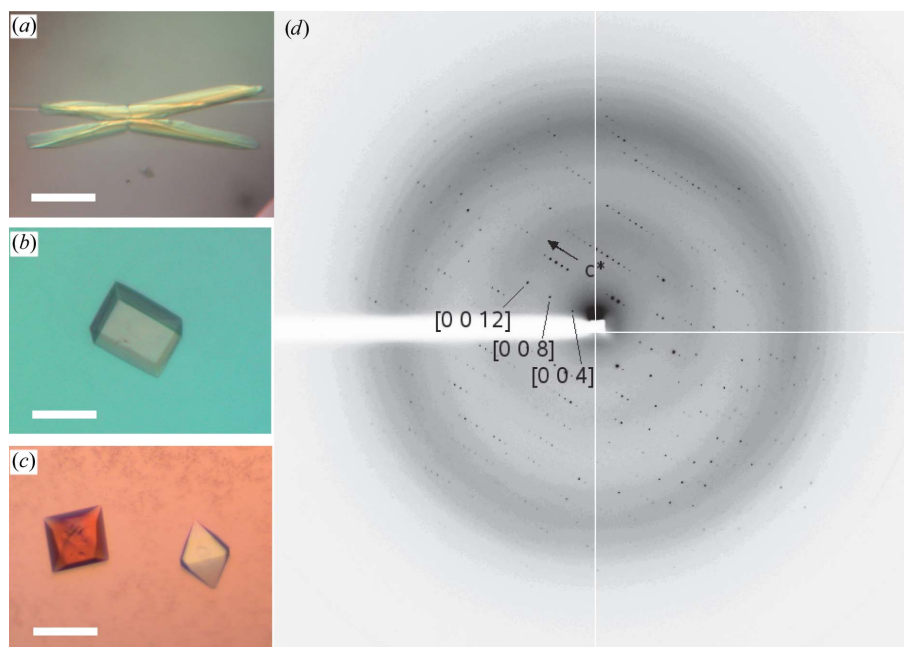
## 2. Material and methods

### 2.1. Cloning, overproduction and protein characterization

Two forms of human mt-TyrRS were produced: the full-length protein and a truncated form lacking the S4-like domain (Fig. 1). Full-length TyrRS was cloned as described in Bonnefond, Fender *et al.* (2005). The truncated version was cloned following the same procedure using the pQE70-mt-TyrRS plasmid as PCR template (primers

were 5'-ATG CGA **GGA TCC** CAC TCG GGC GCT CAG GGG-3' and 5'-GGT GGT **AGA TCT** GTG ATA AAG GGC TTG TGT ACA CC-3', with *Bam*HI and *Bgl*II restriction sites in bold). The cloning vector introduced a tetrapeptide (MRGS) and a His-tagged octapeptide (RSH<sub>6</sub>) at the N- and C-terminal extremities of the cloned TyrRSs, respectively. The overproduction and purification procedures were similar for both proteins and were adapted with minor modifications from Bonnefond, Fender *et al.* (2005). In brief, top10 *Escherichia coli* strains containing either pQE70-mt-TyrRS or pQE70-mt-TyrRS- $\Delta$ S4 were grown in LB medium at 310 K until an OD<sub>600</sub> of 0.7 was reached. Protein expressions were induced overnight at 298 K with 500  $\mu$ M IPTG. After centrifugation, cell pellets were suspended in 25 ml buffer A [50 mM NaH<sub>2</sub>PO<sub>4</sub> pH 8.0, 300 mM NaCl, 5 mM dithioerythritol (DTE)] and sonicated six times for 45 s on ice. Lysed cells were centrifuged at 35 000g for 30 min at 277 K and the supernatants were loaded onto a 2 ml Ni-NTA column (Qiagen). Proteins were eluted with an imidazole gradient (20–500 mM). The pooled enzyme-containing fractions were dialyzed and concentrated for 3 h at 277 K against buffer B (50 mM HEPES–NaOH pH 6.7, 300 mM NaCl, 10 mM DTE) with 50% glycerol. Two further dialyses for 2 h against buffer B with 10% glycerol and then overnight against buffer B gently removed glycerol with minimal volume change. About 5 mg (at  $\sim$ 2–3 mg ml<sup>-1</sup>) full-length or 25 mg (at  $\sim$ 6 mg ml<sup>-1</sup>) truncated electrophoretically pure human mt-TyrRSs were recovered from 1 l LB medium.

The protein particle weight and polydispersity were determined on the basis of size-exclusion chromatography (SEC) and dynamic light-scattering (DLS) measurements. SEC analyses were performed at 277 K on an analytical 15 ml Bio-Prep SE-100/17 column (Bio-Rad) equilibrated in buffer B. The TyrRS elution volume was compared with those of proteins of known molecular weights and hydrodynamic radius. DLS analyses were performed at 293 K on 12  $\mu$ l protein samples at appropriate concentration using a DynaPro DP801



**Figure 2**

Crystals and diffraction properties of human mt-TyrRS. (a) Needles of full-length mt-TyrRS grown in 2 M ammonium sulfate pH 6.5, 10% glycerol, 0.1 M magnesium sulfate. (b) Prismatic crystals of truncated mt-TyrRS- $\Delta$ S4 in 18% (w/v) PEG 4000, 120 mM ammonium acetate, 0.1 M Tris–sodium citrate pH 5.6. (c) Bipyramidal crystals of mt-TyrRS- $\Delta$ S4 in 30% (w/v) PEG 4000, 0.2 M ammonium acetate, 0.1 M sodium acetate pH 4.6. The bars in (a), (b) and (c) correspond to 100  $\mu$ m. (d) Diffraction pattern of a bipyramidal crystal. The image consists of a 0.5° oscillation. The resolution at the corner and the edge of the ADSC Quantum Q210 detector is 2.0 and 2.6 Å, respectively. Reflections of the [00] row (see indices) indicate the presence of a helicoidal fourfold axis along the  $c/c^*$  direction.

instrument (Protein Solutions Inc.). Diffusion coefficients, particle radii and weights were corrected for buffer viscosity and refractive index. Other protein-characterization techniques (SDS-PAGE and N-terminal sequencing) were performed using standard methods. TyrRS activities were determined as described in Bonnefond, Fender *et al.* (2005).

## 2.2. Protein crystallization and crystal analysis

Crystallization trials were set up at 293 K in 96-well plates (Greiner) using a Mosquito robot (TTP LabTech Ltd). Sitting drops of 600 nl–2  $\mu$ l (1:1 to 4:1 mixtures of 3–6 mg ml<sup>-1</sup> protein in buffer *B* and reservoir solutions) were equilibrated by vapour diffusion against 80  $\mu$ l reservoir solution. Prefilled screens (Qiagen) were employed with both full-length and truncated mt-TyrRS. Crystals of either enzyme form grew in about two weeks. They were mounted in cryoloops (Hampton Research), their mother liquor replaced with Paratone (Hampton Research) and flash-frozen in a nitrogen stream. Two data sets for truncated TyrRS were successfully collected to 3.2 and 2.7 Å resolution at 93 K on the ID14-EH4 or ID14-EH1 beamlines at ESRF, France equipped with an ADSC Quantum 315 or 210 CCD detectors, respectively. Data were processed with the *HKL* package (Otwinowski & Minor, 1997).

## 3. Results and discussion

### 3.1. Purification, stability and crystallization of full-length human mt-TyrRS

Purification of full-length human mt-TyrRS (103 kDa, 2 × 458 residues) resulted in pure and active protein (Bonnefond, Fender *et al.*, 2005). Migration of the pure protein on an SDS gel under reducing conditions with DTE in the loading buffer revealed a single band corresponding to the intact monomer, as confirmed by N-terminal sequencing. DLS measurements indicated the protein to be polydisperse and gave a mean hydrodynamic radius of 54 Å, in agreement with the expectation that the enzyme forms an elongated dimer. This enzyme has a tendency to aggregate and to lose its activity at concentrations above 2 mg ml<sup>-1</sup>.

Extensive crystallization assays, conducted with fresh protein samples at 2 mg ml<sup>-1</sup>, led to needle-like crystals (Fig. 2*a*) that diffracted X-rays poorly. Their quality could not be improved. Further assays with the enzyme complexed to *E. coli* tRNA<sup>Tyr</sup> (as verified by DLS analysis) were also ineffective in producing crystals suitable for structure determination. Similar observations were reported for the mitochondrial TyrRS from *N. crassa* (Paukstelis *et al.*, 2005).

### 3.2. Towards high-quality crystals

As was performed with *E. coli* (Kobayashi *et al.*, 2005), *Staphylococcus aureus* (Qiu *et al.*, 2001) and mitochondrial *N. crassa* TyrRSs (Paukstelis *et al.*, 2005), the C-terminal domain of human mt-TyrRS, which is likely to be floppy, was removed. This domain, which is analogous to ribosomal protein S4, was delimited by multiple sequence alignment with bacterial TyrRSs and assigned as the 102 last residues (Fig. 1). The truncated version therefore encompasses the N-terminal catalytic domain and the anticodon-binding domain.

The purified truncated TyrRS migrates as a unique 40 kDa band on SDS gel, activates tyrosine and charges *E. coli* tRNA<sup>Tyr</sup>. This protein (2 × 356 residues) is monodisperse in DLS, with a mean hydrodynamic radius of 38 Å, in agreement with a homodimeric structure. This conclusion was confirmed by SEC analyses. Importantly, dele-

**Table 1**

X-ray analysis of a human mt-TyrRS- $\Delta$ S4 crystal.

Data were collected from a bipyramidal crystal (see Fig. 2*c*); values in parentheses are for the highest resolution shell.

Beamline	ID14-1, ESRF
Wavelength (Å)	0.934
Space group	<i>P</i> 4 <sub>3</sub> 2 <sub>1</sub> 2
Unit-cell parameters (Å)	<i>a</i> = 78.8, <i>c</i> = 121.1
Crystal mosaicity (°)	0.7
Resolution range (Å)	2.7–30 (2.7–2.8)
No. of observations	534152
No. of unique reflections	11044
Completeness (%)	99.9 (100)
Multiplicity	13.1 (13.3)
<i>R</i> <sub>merge</sub> † (%)	6.8 (37.2)
<i>I</i> / <i>σ</i> ( <i>I</i> )	27.8 (4.9)
Matthews coefficient (Å <sup>3</sup> Da <sup>-1</sup> )	2.4
Solvent content (%)	48
Asymmetric unit content	1 monomer

$$\dagger R_{\text{merge}} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$$

tion of the S4-like domain significantly increases the solubility of the synthetase, which can now be concentrated up to 20 mg ml<sup>-1</sup>, thus opening new crystallization possibilities.

More than 600 different crystallization conditions were assayed with human mt-TyrRS- $\Delta$ S4, of which 40 gave crystals. Most of them contained PEGs as the crystallizing agent. The largest crystals display prismatic and bipyramidal morphologies (Figs. 2*b* and 2*c*). The better diffracting bipyramidal crystals (Fig. 2*c*) grew reproducibly at pH 4.6 and in the presence of 30% (*w/v*) PEG 4000. By changing the protein:precipitant ratio to 2:1, their dimensions were enlarged and their diffraction limit extended from 3.2 to 2.7 Å.

### 3.3. Preliminary X-ray data

Crystals of truncated TyrRS were analyzed using synchrotron radiation. Statistics and other crystallographic data for the best crystal are given in Table 1. Based on systematically absent reflections, the space group is either *P*4<sub>3</sub>2<sub>1</sub>2 or *P*4<sub>1</sub>2<sub>1</sub>2 (Fig. 2*d*). Analysis of the solvent content performed with the *CCP4* package (Collaborative Computational Project, Number 4, 1994) gave a unique solution consistent with one polypeptide chain (356 residues) per asymmetric unit. Molecular-replacement (MR) trials were performed with the *CaspR* MR webservice (Claude *et al.*, 2004) using the structure of *E. coli* TyrRS (PDB code 1wq3) and derived search models. In the resolution range 3–15 Å, space group *P*4<sub>3</sub>2<sub>1</sub>2 gave a better solution (correlation = 52%, *R* factor = 53%) than its enantiomorph *P*4<sub>1</sub>2<sub>1</sub>2 (correlation = 46%, *R* factor = 56%). Notice that the subunits of the biological dimer are related by the crystallographic twofold symmetry axis. After rigid-body refinement, the MR solution led to an interpretable electron-density map. Refinement of the model is in progress.

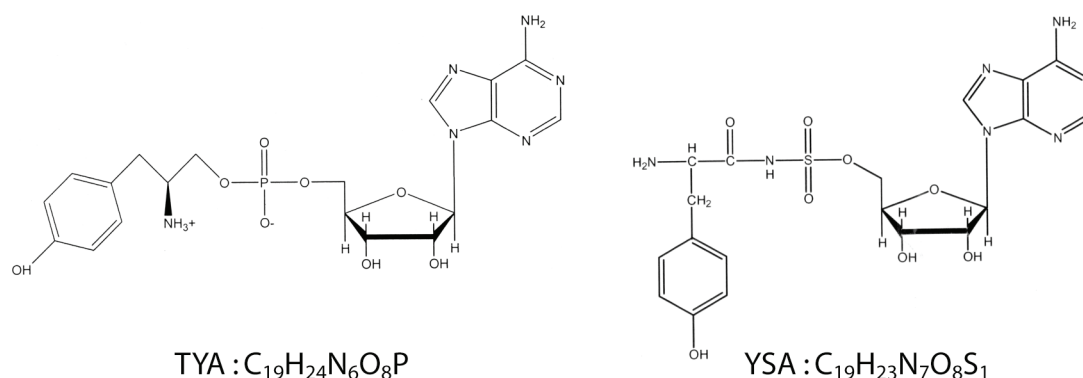
The authors acknowledge the teams at the ID14 beamlines at ESRF (Grenoble, France) for assistance during data collection as well as Guillaume Bec for help with the nanodrop crystallization robot and Marie Sissler for stimulating discussions. This work was supported by the Centre National de la Recherche Scientifique, Université Louis Pasteur (Strasbourg) and by grants from the French Ministry for Research (ACI-BCMS 042358). LB was supported by a grant from the French Ministry for Research and CS was the recipient of a Marie Curie European Reintegration Grant (MERG-CT-2004-004898).

## References

- Abergel, C., Chenivesse, S., Byrne, D., Suhre, K., Arondel, V. & Claverie, J.-M. (2005). *Acta Cryst.* **F61**, 212–215.
- Bedouelle, H. (2005). *Aminoacyl-tRNA Synthetases*, edited by M. Ibba, C. Francklyn & S. Cusack, pp. 111–124. Georgetown, TX, USA: Landes Biosciences.
- Bhat, T. N., Blow, D. M., Brick, P. & Nyborg, J. (1982). *J. Mol. Biol.* **158**, 699–709.
- Bonnefond, L., Fender, A., Rudinger-Thirion, J., Giegé, R., Florentz, C. & Sissler, M. (2005). *Biochemistry*, **44**, 4805–4816.
- Bonnefond, L., Frugier, M., Giegé, R. & Rudinger-Thirion, J. (2005). *RNA*, **11**, 558–562.
- Chimnarok, S., Jeppesen, M. G., Shimada, N., Suzuki, T., Nyborg, J. & Watanabe, K. (2004). *Acta Cryst.* **D60**, 1319–1322.
- Claude, J.-B., Suhre, K., Notredame, C., Claverie, J.-M. & Abergel, C. (2004). *Nucleic Acids Res.* **32**, W606–W609.
- Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* **D50**, 760–763.
- Cusack, S., Berthet-Colominas, C., Härtlein, M., Nassar, N. & Leberman, R. (1990). *Nature (London)*, **347**, 249–255.
- Eriani, G., Delarue, M., Poch, O., Gangloff, J. & Moras, D. (1990). *Nature (London)*, **347**, 203–206.
- Florentz, C., Sohm, B., Tryoen-Tóth, P., Pütz, L. & Sissler, M. (2003). *Cell. Mol. Life Sci.* **60**, 1356–1375.
- Giegé, R., Sissler, M. & Florentz, C. (1998). *Nucleic Acids Res.* **26**, 5017–5035.
- Jacobs, H. & Turnbull, D. (2005). *Trends Genet.* **21**, 312–314.
- Kobayashi, T., Takimura, T., Sekine, R., Vincent, K., Kamata, K., Sakamoto, K., Nishimura, S. & Yokoyama, S. (2005). *J. Mol. Biol.* **346**, 105–117.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Paukstelis, P. J., Coon, R., Madabusi, L., Nowakowski, J., Monzingo, A., Robertus, J. & Lambowitz, A. M. (2005). *Mol. Cell*, **17**, 417–428.
- Qiu, X., Janson, C. A., Smith, W. W., Green, S. M., McDevitt, P., Johanson, K., Carter, P., Hibbs, M., Lewis, C., Chalker, A., Fosberry, A., Lalonde, J., Berge, J., Brown, P., Houge-Frydrych, C. S. V. & Jarvest, R. L. (2001). *Protein Sci.* **10**, 2008–2016.
- Sissler, M., Pütz, J., Fasiolo, F. & Florentz, C. (2005). *Aminoacyl-tRNA Synthetases*, edited by M. Ibba, C. Francklyn & S. Cusack, pp. 271–284. Georgetown, TX, USA: Landes Biosciences.
- Yaremchuk, A., Kriklivyi, I., Tukalo, M. & Cusack, S. (2002). *EMBO J.* **21**, 3829–3840.

## 2. La structure de la TyrRS mitochondriale

Une première carte de densité a été calculée par remplacement moléculaire grâce aux jeux de données de la mt-TyrRS tronquée enregistrés à la résolution de 2.7 Å. Cette carte a permis de construire la structure du domaine catalytique à l'exception de quelques secteurs désordonnés tels que la boucle du "KMSKS". Par contre le domaine de liaison à l'anticodon était très mal défini et les tentatives de reconstruction hasardeuses. Pour essayer d'améliorer la qualité des cristaux et d'obtenir des données de diffraction de meilleure qualité nous avons réalisé des essais de co-cristallisation avec les différents substrats de la mt-TyrRS : ARNt, tyrosine, ATP et analogues de l'adénylate de tyrosine, seuls ou combinés. Des cristaux ont rapidement été obtenus en présence de deux analogues d'adénylate de tyrosine.

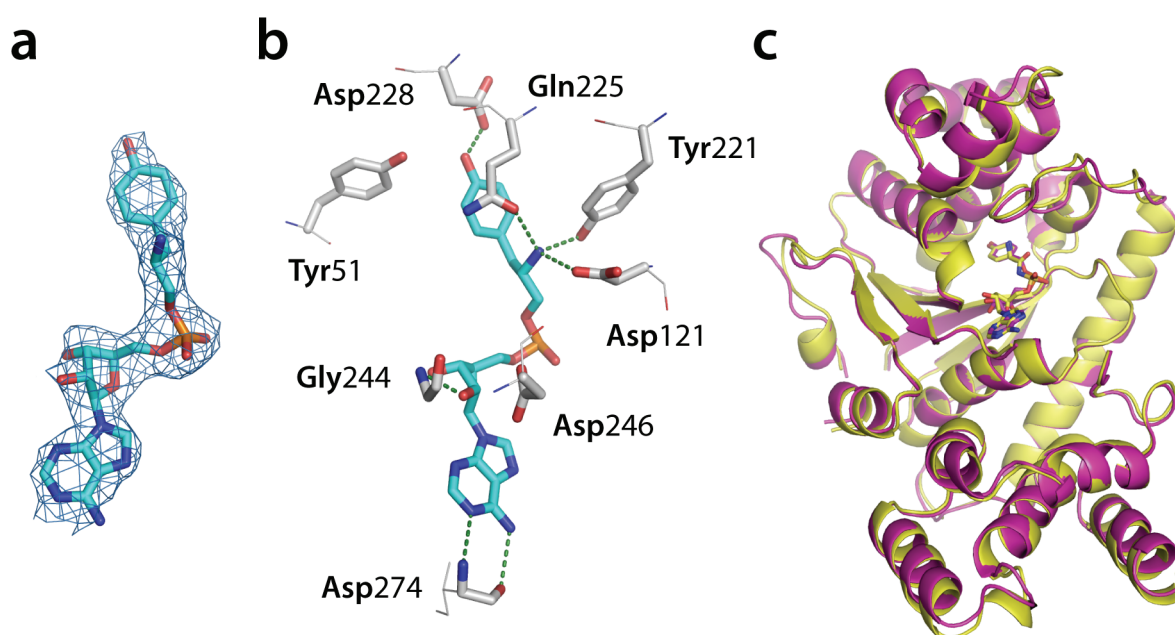


**Figure 20. Formules des dérivés d'adénylate de tyrosine utilisés en cristallographie**

Le premier analogue est le 5'-O-[N-(L-tyrosyl)sulfamoyl]adénosine aussi appelé YSA (ou Tyr-AMS dans l'**article 5**) dans lequel le phosphate a été substitué par un soufre non hydrolysable par la TyrRS (synthétisé par RNATUK, USA, TX). Le second analogue est l'acide phosphorique 2-amino-3-(4-hydroxy-phenyl)-prolyl ester adénosine-5'yl ester, abrégé TYA (synthétisé par l'équipe du Dr. Chenevert à Québec, à l'origine de nombreux analogues d'aminoacyl-adénylates; Bernier et al., 2000) (**figure 20**). Bien que présentant la même limite de diffraction vers 2.7 Å que les cristaux de l'enzyme libre (**tableau 8**), les cartes de densité calculées à partir de ces nouveaux jeux de données étaient de meilleure qualité et ont permis de reconstruire le domaine de liaison à l'anticodon dans son intégralité.

Enzyme	mt-TyrRS-ΔS4	mt-TyrRS-ΔS4	mt-TyrRS-ΔS4
Ligand	-	YSA	TYA
Ligne synchrotron	ID14-1 (ESRF)	ID23-2	ID23-2 (ESRF)
Longueur d'onde (Å)	0,93	0,87	0,87
Groupe d'espace	$P4_32_12_1$	$C22_1$	$P4_32_12$
Paramètres de la maille a/b/c (Å)	78,8 / 78,8 / 121,1	110,7 / 114,2 / 120,2	79,5 / 79,5 / 120,2
Mosaïcité (deg)	0,66	0,66	0,28
Résolution (Å)	2,7–30 (2,7–2,8)	2,7–30 (2,7–2,8)	2,8–30 (2,8–2,9)
Nombre de réflexions observées	534152	302021	327050
Nombre de réflexions uniques	11044	21068	10029
Complétude (%)	99,9 (100)	99,8 (99,7)	99,9 (100)
Multiplicité	13,1 (13,3)	3,9 (3,8)	7,9 (8,2)
Rmerge (%)	6,8 (37,4)	8,4 (48,6)	8,6 (46)
I/s (I)	28 (5)	19 (4,7)	17 (4,5)
Proportion de solvant (%)	48	48	48
Contenu de l'unité asymétrique	1 monomère	2 monomères	1 monomère

**Tableau 6. Jeux de données pour les formes libre et complexées de la mt-TyrRS-ΔS4**  
Les valeurs entre parenthèses correspondent à la tranche de haute résolution.

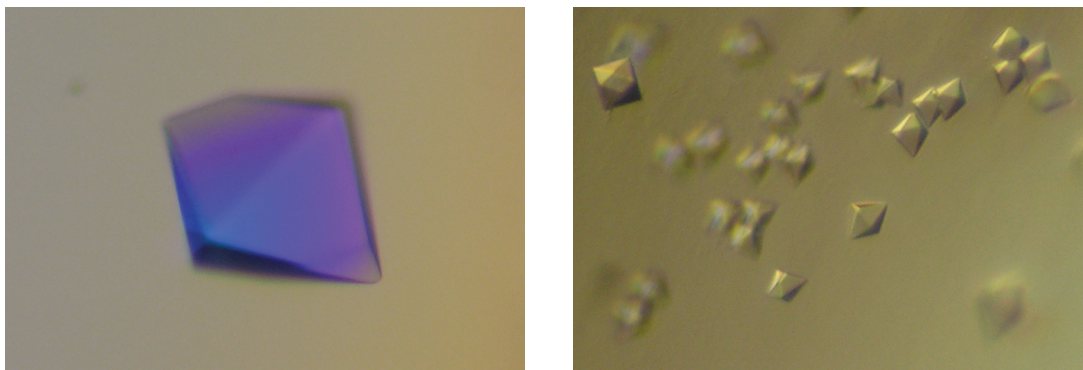


**Figure 21. Structure de la mt-TyrRS-ΔS4 complexée à l'analogue de tyrosyl-adénylate TYA**  
(a) Densité électronique autour du TYA. La carte est contournée à 1.4  $\sigma$ . (b) Organisation du site catalytique avec les résidus qui contactent le substrat *via* des liaisons hydrogène (en vert). Les résidus Tyr51 et Asp246 sont à plus de 3 Å du substrat, mais ont été représentés car ils contactent le substrat dans la structure de l'enzyme complexée au YSA (voir l'article ci-après). (c) Superposition d'un monomère de TyrRS/TYA (en violet) sur le monomère A de la TyrRS/YSA (en jaune). L'analogue d'adénylate de tyrosine lié dans le site actif de chaque complexe est représenté.



Bien que l'affinement ne soit pas terminé ( $R=30$ ,  $R_{\text{libre}}=35$ ), la structure de la mt-TyrRS- $\Delta S4$  dans laquelle le substrat TYA est très bien défini (**figure 21.a**) permet déjà d'observer les résidus du site actif en interaction avec l'analogue de tyrosyl-adénylate (**figure 21.b**). La superposition de cette structure avec celle obtenue avec le second analogue est presque parfaite, aussi bien au niveau de l'enzyme que du ligand (**figure 21.c**). La structure de la mt-TyrRS- $\Delta S4$  complexée au YSA obtenue à 2.7 Å n'a pas été affinée car des données à plus haute résolution ont été obtenues. La description de cette structure et la discussion de ses différentes caractéristiques sont développées dans l'article qui suit (**article 5**).

Nous avons aussi préparé des cristaux de la protéine marquée au sélénium (**figure 22**) et enregistré des jeux de données pour pouvoir calculer une carte expérimentale et ainsi se défaire du biais introduit par le modèle employé dans la technique de remplacement moléculaire. Malheureusement, l'intensité du signal anomal est trop faible pour nous permettre de phaser les données.



**Figure 22. Photos de cristaux de mt-TyrRS- $\Delta S4$  marquée au sélénium**

Les cristaux sont obtenus en présence de YSA (5'-O-[N-(L-tyrosyl)sulfamoyl]adénosine) et ont une taille maximale de 100 micromètres.

Parallèlement nous avons optimisé la qualité des cristaux de l'enzyme native complexée au YSA en utilisant une gamme d'additifs et de tampons autour des conditions préalablement identifiées. Les cristaux obtenus présentent une autre forme cristalline et diffractent les rayons X à 2.2 Å de résolution. La structure reconstruite à partir de ce jeu de données a été étudiée en détail donnant lieu à une publication soumise à la revue "Structure" en Avril 2007.

## **Propriétés générales et particularités de la structure de la TyrRS mitochondriale humaine**

Luc Bonnefond, Magali Frugier, Elodie Touzé, Bernard Lorber, Catherine Florentz, Richard Giegé, Claude Sauter et Joëlle Rudinger-Thirion

**2007**

Structure (soumis le 20 Avril 2007)

En dépit des nombreuses études réalisées sur les aminoacyl-ARNt synthétases, les connaissances concernant les synthétases mitochondriales sont encore limitées. Dans cet article nous présentons la première structure d'une synthétase mitochondriale humaine, la tyrosyl-ARNt synthétase (mt-TyrRS) dépourvue de son domaine C-terminal S4-like, à 2,2 Å de résolution. Cette mini-TyrRS conserve son activité enzymatique et ressemble aux TyrRS eubactériennes dépourvues de leur domaine C-terminal S4-like avec un site de fixation de la tyrosine strictement conservé et des résidus liant l'adénylate typiques des synthétases de classe I. Deux insertions, présentes à sa surface, correspondent à des séquences conservées chez les TyrRS mitochondriales. Elle diffère des autres TyrRS connues par son potentiel électrostatique de surface, notamment la TyrRS mitochondriale de *Neurospora crassa* et son homologue cytoplasmique humain. La mt-TyrRS humaine homodimérique présente une asymétrie qui se propage de l'interface de dimérisation vers les deux sites catalytiques et jusqu'aux extrémités. Enfin, la mutagénèse réalisée sur le domaine catalytique démontre l'importance du résidu Ser200 pour l'aminoylation en accord avec l'implication de la base discriminatrice A73 plutôt que la paire G1-C72 dans l'identité tyrosine mitochondriale.

# **Crystal Structure of Human Mitochondrial Tyrosyl-tRNA Synthetase Reveals Common and Idiosyncratic Features**

Luc Bonnefond, Magali Frugier, Elodie Touzé, Bernard Lorber, Catherine Florentz, Richard Giegé\*, Claude Sauter, and Joëlle Rudinger-Thirion

Département «Machineries Traductionnelles», Architecture et réactivité de l'ARN, Université Louis Pasteur de Strasbourg, CNRS, IBMC, 15 rue René Descartes, 67084 Strasbourg, France.

Running Title: **Crystal Structure of Human Mitochondrial TyrRS**

Correspondence: [r.giege@ibmc.u-strasbg.fr](mailto:r.giege@ibmc.u-strasbg.fr);

Phone: 33 (0)3 88 41 70 58 ; Fax: 33 (0)3 88 60 22 18

## **SUMMARY**

We report the first structure of a human mitochondrial synthetase, namely tyrosyl-tRNA synthetase (mt-TyrRS) in complex with an adenylate analog at 2.2 Å resolution. The structure is that of an active enzyme deprived of the C-terminal S4-like domain and resembles eubacterial TyrRSs with a canonical tyrosine-binding pocket and adenylate-binding residues typical of class I synthetases. Two bulges at the enzyme surface, not seen in eubacterial TyrRSs, correspond to conserved sequences in mt-TyrRSs. The synthetase electrostatic surface potential differs from that of other TyrRSs, including the human cytoplasmic homolog and the mitochondrial one from *Neurospora crassa*. The homodimeric human mt-TyrRS shows an asymmetry propagating from the dimer interface towards the two catalytic sites and extremities of each subunit. Mutagenesis of the catalytic domain reveals functional importance of Ser200 in line with an involvement of A73 rather than N1–N72 in tyrosine identity.



## INTRODUCTION

Among aminoacyl-tRNA synthetases, tyrosyl-tRNA synthetases (TyrRSs) present unique features. Although they belong to class I synthetases with the master signature sequences "HIGH" and "KMSKS" and a Rossmann-fold catalytic domain (reviewed by Bedouelle, 2005; Bonnefond et al., 2005c), they are homodimers and recognize tRNA from the major groove side of the amino acid acceptor stem, in a way reminiscent to what found for class II synthetases (Bedouelle and Winter, 1986; Lee and RajBhandary, 1991; Yaremchuk et al., 2002). They are also peculiar with regard to the tRNA identity rules, with the G1–C72 base-pair contributing to tyrosine identity in all eubacterial tRNA<sup>Tyr</sup> (Quinn et al., 1995) and the reverse C1–G72 pair in archaeal/eukaryal species, where it is the major tyrosine determinant (Lee and RajBhandary, 1991; Quinn et al., 1995; Fechter et al., 2000; 2001). This peculiarity contrasts with what observed with most identities that are generally conserved in evolution (Giegé et al., 1998; Beuning and Musier-Forsyth, 1999). Further, it explains why cross-tyrosylations between tRNA and TyrRSs originating from different kingdoms of life do not occur (Kleeman et al., 1997; Wakasugi et al., 1998). The structures of TyrRSs present a conserved gross organization with a N-terminal catalytic domain followed by the C-terminal anticodon-binding region. Interestingly TyrRSs show large variations in sequence and length, due to insertions and appended domains (Wolf et al., 1999; Bedouelle, 2005). Additional functions result from these differences, like involvement in splicing for mitochondrial *Neurospora crassa* TyrRS (Cherniak et al., 1990) and cytokine activity for a fragment of human cytoplasmic TyrRS (Wakasugi and Schimmel, 1999).

The present contribution focuses on human mitochondrial TyrRS (mt-TyrRS). Although eubacterial-like in terms of overall organization, human mt-TyrRS, unexpectedly charges as well tRNA<sup>Tyr</sup> species from eubacteria with a G1–C72 pair as from eukarya with C1–G72. Thus, tyrosylation specificity of human mt-TyrRS is unresponsive to the N1–N72 pair, making this protein the first TyrRS with non-conservation of a universal tyrosine identity position (Bonnefond et al., 2005b). To gain a deeper molecular understanding of this surprising property, knowledge on the three-dimensional structure is important. But this knowledge is lacking so far, since no crystallographic structure of any such human enzyme (except SerRS from closely related bovine mitochondria, Chimnaronk et al., 2005) is presently available. Also, many human pathologies have been correlated with defects in mitochondrial tRNAs (Florentz et al., 2003; Taylor and Turnbull, 2005; Brandon et al., 2006) and are expanding towards components of the mitochondrial translation machinery (Jacobs and Turnbull, 2005; Antonicka et al., 2006; Scheper et al., 2007), appealing for structural knowledge on these macromolecules. From another viewpoint, the tyrosine system with known TyrRS structures

from all kingdoms of life, represents a robust model to understand the evolution of aminoacyl-tRNA synthetases. Here we present the structure and functional features of a C-terminally truncated but active version of this human TyrRS in complex with a tyrosyl-adenylate analog. Mitochondrial characteristics, structural asymmetry and idiosyncratic amino acids in the tRNA acceptor stem binding domain of the homodimeric enzyme will be discussed.

## RESULTS

### Solution and Crystallographic Properties of Human mt-TyrRS

Numerous attempts to crystallize native human mt-TyrRS led to crystals not suitable for structure resolution (Bonfond et al., 2007). Based on the experience on *Staphylococcus aureus*, *Escherichia coli* and mitochondrial *N. crassa* TyrRSs (Qiu et al., 2001; Kobayashi et al., 2005; Paukstelis et al., 2005), the potentially floppy C-terminal domain, analogous to ribosomal protein S4 (Figure 1A,B), was removed. The truncated TyrRS (2x356 residues), called mt-TyrRS- $\Delta$ S4, consists therefore in the N-terminal catalytic domain followed by the helical  $\alpha$ -ACB sub-structure of the anticodon-binding region. When purified, this protein is monodisperse and has a  $\sim 38$  Å hydrodynamic radius (Bonfond et al., 2007). It is active for tyrosine activation and charging of native *E. coli* tRNA<sup>Tyr</sup>, although tyrosylation efficiency  $k_{\text{cat}} / K_M$  is decreased  $\sim 100$ -fold mainly due to a decreased  $k_{\text{cat}}$ . Notice the inability of mt-TyrRS- $\Delta$ S4 to charge other native (yeast) or transcribed (*E. coli*, yeast, human cytosolic and mitochondrial) tRNA<sup>Tyr</sup> species as opposed to the native mt-TyrRS (Bonfond et al., 2005b). Interestingly, deletion of the S4-like domain significantly increases the solubility of the synthetase, as compared to the full-length protein. This property is favorable for crystallization and, indeed tetragonal crystals diffracting to 2.7 Å resolution could be grown (Bonfond et al., 2007). The resolution of the latter could be extended to 2.2 Å when crystallization was conducted in the presence of the tyrosyl-adenylate analog 5' *O*-[*N*-(L-tyrosyl) sulfamoyl] adenosine (Tyr-AMS). We note that enzymatic activity and crystallization necessitate reducing conditions, as was the case for bovine mt-SerRS (Chimnaronk et al., 2004).

### Overall Structure of Human mt-TyrRS- $\Delta$ S4

The structure of human mt-TyrRS- $\Delta$ S4 in complex with Tyr-AMS was solved by molecular-replacement (MR) using the *E. coli* TyrRS (PDB code 1VBM) (Kobayashi et al., 2005) as initial search model. The refined model at 2.2 Å resolution, yields a crystallographic *R*-factor of

19.6% with a free *R*-factor of 24.4% (Table 1). It is based on remarkably good electron density (see the Supplemental Data 1). A second structure at 2.7 Å resolution corresponds to the free enzyme but gives only a partial view of the molecule. The C-terminal domain is globally disordered, whereas the catalytic site is well defined. In what follows we will concentrate on the description of the structure including the adenylate analog.

The dimeric structure has an elongated shape with the C-terminal ends in distal location (110 Å in length) (Figure 2). The structure reveals the modular architecture of the synthetase with the two N-terminal catalytic domains in central location of the protein. These catalytic domains encompass the Rossmann-folds with their typical parallel β-strands (β2 to β6) and the two active sites filled with Tyr-AMS in extended conformation. The anticodon-binding modules, constituted by five α-helices (α11 to α15), are located at both distal extremities of the dimer. Finally the interface between the two subunits is made by the CP1 domains (Figure 2). Among the two class I synthetase signature sequences, only HVGH is seen. This motif is close to the adenosine moiety of bound Tyr-AMS. The KLGKS signature sequence is located in the flexible loop between β6 and α11 connecting the catalytic to the anticodon binding domains. Notice the location of two helical structures close to the dimerization interface, named cluster 1 (α7 and α8) and cluster 2 (α10) (Figures 1B and 2), that were shown to be essential for species-specific recognition of the amino acid acceptor stem of tRNA<sup>Tyr</sup> (Nair et al., 1997).

In contrast to the crystals of free enzyme that contain one monomer in the asymmetric unit (Bonfond et al., 2007), those of the complex with Tyr-AMS contain the whole homodimer (Table 1). Remarkably, the two monomers are not identical in the crystals, as revealed by an overall root-mean-square deviation (r.m.s.d.) of 0.5 Å, which can reach 3.5 Å in several parts of the synthetase subunits (Figure 3; see the Supplemental Data 2). Most prominent structural changes occur in three regions of the subunits. First in the CP1 region making the interface between the two subunits (in helix α8 and the following loop connecting to helix α9), second in the “ins1” region of the catalytic domain (see below) and third in the α-ACB region that moves as a rigid body (Figure 3).

### **Comparison with other TyrRSs**

TyrRS structures originating from 11 organisms, namely four Archaea (Zhang et al., 2005; Kuratani et al., 2006), four Eubacteria (Brick et al., 1989; Qiu et al., 2001; Kobayashi et al., 2005; Yaremchuk et al., 2005), an Eukarya (Yang et al., 2002), a virus (Abergel et al., 2005) and the *N. crassa* mitochondria (Paukstellis et al., 2005), have already been described. Here we add the structure from human mitochondria. Most of these structures, including human

mt-TyrRS, are deprived of the C-terminal domain because its structural flexibility hampers crystallization. Although these enzymes present common characteristics, as an overall similar shape, they show large variations in their electrostatic surface potentials as depicted in Figure 4 for five representative structures. These variations concern both the TyrRS face recognizing the tRNA and the opposite face that does not contact tRNA. Below we further compare selected structural features of human mt-TyrRS with those of known TyrRSs including its cytoplasmic human homolog.

Sequence alignments show high similarities between human mt-TyrRS and eubacterial TyrRSs (e.g. 37%, 35% and 35% identity with native TyrRSs from *Bacillus stearothermophilus*, *E. coli* and *S. aureus*, respectively, despite lower homology in the anticodon-binding domain) (Bonfond et al., 2005a). Further, the overall architecture of mt-TyrRS- $\Delta$ S4 as well is similar to that of known eubacterial TyrRSs, with r.m.s.d. of 1.30 Å for *E. coli* (PDB code 1VBM), 1.33 Å for *B. stearothermophilus* (PDB code 3TS1) and 1.36 Å for *S. aureus* (PDB code 1JIL) TyrRSs. Figure 5A compares human mt-TyrRS- $\Delta$ S4 and *E. coli* (PDB code 1VBM) TyrRS, its closest relative, in complex with adenylate. Despite the similarities, one notices two structural idiosyncrasies in the mitochondrial structure. They are an enlarged loop located between  $\beta$ 1 and  $\beta$ 2 strands ("ins1", residues 56–70) and a lengthened  $\alpha$ 4-helix ("ins2", residues 152–166) in the catalytic domain (Figures 5A). Inspection of mt-TyrRS sequences shows presence of the "ins2"-helix in most organisms, but with changes in length and sequence, and highlights the vertebrate character of the "ins1"-loop (see the Supplemental Data 3). Gross docking of *T. thermophilus* tRNA<sup>Tyr</sup> on human mt-TyrRS suggests that the "ins1"-loop contacts the ribose-phosphate backbone of the tRNA acceptor helix and consequently may contribute to the tRNA anchoring onto the synthetase.

To better understand mitochondrial idiosyncrasies, we compare the structure of human mt-TyrRS with that from the yeast *N. crassa* mitochondria, known to participate in splicing of mt-group I introns (Akins and Lambowitz, 1987). Overall, the two structures are similar throughout their whole sequence with an r.m.s.d. of 1.75 Å (Figure 5B,C). Three insertions (H0, I and II) in the catalytic domain of the *N. crassa* enzyme were proposed to be involved in splicing (Paukstelis et al., 2005). Interestingly, the splicing platform is opposite to the tRNA recognizing face and has a positive electrostatic surface potential (Figure 4). Out of the three insertions, only insertion I is found in human mt-TyrRS, where it corresponds to "ins2", with however fewer residues (Figures 1B and 5B,C).

Finally it is worth to compare the crystallographic structure of human mt-TyrRS with that of its cytosolic counterpart. This enzyme, as well, is merely known under a minimalist form (Yang et al., 2002), lacking the eukaryal-specific EMAP-like Ct-terminal domain solved

independently (Yang et al., 2003). Although both mitochondrial and cytosolic TyrRSs have the same overall shape, they strikingly differ in the surface distribution of the electrostatic potential (Figure 4). This might correlate either with a functioning of the two human enzymes in different physico-chemical environments or with the binding of alternate partners in cytoplasm and mitochondria. Note that the ELR tripeptide responsible for the IL8 cytokine activity of cytoplasmic mini-TyrRS (Yang et al., 2002) is not present in the mitochondrial enzyme.

### **Snapshots in the Active Site**

The difference in resolution of free (the original crystal form) and Tyr-AMS bound mt-TyrRS- $\Delta$ S4 forms (2.7 and 2.2 Å, respectively) suggests a stabilization of the enzyme conformation after binding the adenylate analog. The active site is well defined (Figure 6A) in the two enzyme forms, with the exception of the degenerated class I signature KLGKS that is disordered in both structures, as previously found in other TyrRSs (e.g. *B. stearothermophilus*, PDB code 3TS1, Brick et al., 1989, and *E. coli*, PDB code 1VBM, Kobayashi et al., 2005). This signature sequence, located at the periphery of the active site where it plays a functional role in tyrosylation (Xin et al., 2000), becomes ordered in the TyrRS when in complex with ATP or tRNA<sup>Tyr</sup> (Yaremchuk et al., 2002). In the human enzyme, Tyr-AMS is clearly defined in the electron density map (Figure 6A) and accommodates well in the active site pocket (Figure 6B), where eight amino acids make hydrogen bonds with the adenylate analog (Figure 6C). Among them two are specific of class I synthetases (Gly244 and Asp246) and warrant discrimination between ATP and dATP, four are strictly conserved in all TyrRSs (Tyr77, Tyr221, Gln225, Asp228) and are responsible for tyrosine recognition, and the two remaining (Asp121, Ile274) are common to mitochondrial and eubacterial TyrRSs.

### **A Functional View on tRNA<sup>Tyr</sup> Acceptor End Recognition Site on mt-TyrRS**

The active site where catalysis occurs, and defined above by the binding pocket of adenylate, is completed by structural elements that specifically recognize the tyrosine accepting arm of tRNA<sup>Tyr</sup> (Figure 7A). They include two helices of fourteen amino acids each, the so-called clusters 1 and 2 (Nair et al., 1987) (Figure 1B) surrounding the catalytic region. These clusters have the same overall helical geometry than in other known eubacterial TyrRSs, with a marked kink in cluster 1, and superimpose best with *B. stearothermophilus* TyrRS without any distortion of their peptide backbones (r.m.s.d. of 0.9 Å) (Figure 7A). Notice that the two helices ( $\alpha$ 7 and  $\alpha$ 8) constituting cluster 1 are more open in the

mitochondrial enzyme. In *T. thermophilus* TyrRS, either free or in complex with tRNA (Yaremchuk et al., 2002), superimposition is worse (r.m.s.d. of 2.0 Å) because of an intrinsic enlargement of the cleft between the two clusters (up to 2.5 Å) (Figure 7A).

Sequence comparison of TyrRSs (Bonnefond et al., 2005b), mutational studies on the *B. stearothermophilus* enzyme (Labouze and Bedouelle, 1989) and inspection of crystallographic structure of an eubacterial tRNA<sup>Tyr</sup>•TyrRS complex (Yaremchuk et al., 2002) pinpoint three potentially essential amino acids for tRNA<sup>Tyr</sup> recognition in the helical clusters of human mt-TyrRS. These amino acids are close to the tRNA tyrosine identity determinants A73 and G1–C72 as seen in the crystal structure of the *T. thermophilus* tRNA<sup>Tyr</sup>•TyrRS complex (Figure 7B). We have explored these three positions that are Ser200 and Gln202 in cluster 1 and Met252 in cluster 2 (Figure 7A). In mt-TyrRSs, Ser200 is highly conserved but replaced by Glu154 or Ala150 in *T. thermophilus* or *B. stearothermophilus* TyrRSs, respectively. Whereas in the *T. thermophilus* complex, Glu154 is seen to contact directly C72, in the *B. stearothermophilus* enzyme, Ala150 was proposed to recognize discriminator base A73. The second amino acid, Gln202, is specific to human mt-TyrRS and, as suggested by docking tRNA on the synthetase, it may interact with the 3'-sequence of tRNA<sup>Tyr</sup> containing A73 that is the strongest tyrosine identity element in human mt-tRNA<sup>Tyr</sup> (Bonnefond et al., 2005c). The third position, Met252, is strictly conserved in archaeal TyrRSs but is variable in eubacterial TyrRSs (Bonnefond et al., 2005b). This Met residue is homologous to residues interacting with C1 or G1 bases, as seen in two crystal structures (Yaremchuk et al., 2002; Kobayashi et al., 2003).

The three amino acids were mutagenized in full-length human mt-TyrRS and tyrosylation assays were performed with human mt-tRNA<sup>Tyr</sup> transcripts. All three mutants are fully active in ATP/PPi exchange (not shown), indicating that the site for tyrosyl-adenylate formation remains unperturbed upon mutation. The effects of the mutations on tRNA<sup>Tyr</sup> tyrosylation are summarized in Table 2. Surprisingly, mutating Gln202 and Met252 does not show any effect on tRNA<sup>Tyr</sup> charging, except faint compensating effects on  $k_{cat}$  and  $K_M$ . Remarkably, replacing Ser200 by Glu completely abolishes tyrosylation activity.

## DISCUSSION

Structural biology of aminoacyl-tRNA synthetases started in the seventieth with the crystallographic structure of TyrRS from *B. stearothermophilus* (Irwin et al., 1976). Today, TyrRSs are among the best-known synthetases with a panel of documented structures, belonging to the three kingdoms of life, and even from the giant Mimivirus infecting the

protist *Acanthamoeba polyphaga*. However, the structural knowledge of organellar synthetases remains limited. The TyrRS investigated in this work led to the first crystal structure of a human mitochondrial aminoacyl-tRNA synthetase. For crystallogensis reasons, this structure corresponds to a truncated protein lacking the S4-like domain. This enzyme is active for both tyrosine activation and *E. coli* tRNA<sup>Tyr</sup> tyrosylation. Likewise, the structure of another mitochondrial and two eubacterial truncated TyrRSs could be solved (Qiu et al., 2001; Kobayashi et al., 2005; Paukstelis et al., 2005). These mini-TyrRSs are active in tyrosine activation, but in contrast to human mt-TyrRS- $\Delta$ S4, the truncated *E. coli* enzyme does not charge tRNA (Kobayashi et al., 2005). In the case of the *S. aureus* or *N. crassa* mitochondria enzymes, aminoacylation activity was not tested (Qiu et al., 2001) or was considered *a priori* not possible because part of the anticodon binding structure, namely the S4-like domain, is lacking (Paukstelis et al., 2005). Note that the S4-like domain interacts also with the large variable region of eubacterial tRNA<sup>Tyr</sup> (Yaremchuk et al., 2002) that is absent in all mt-tRNA<sup>Tyr</sup> species (Sprinzl and Vassilenko, 2005). Thus, one can question why evolution has maintained the S4-like module in mitochondria. We suggest that the appended S4-like domain reinforces the interaction between mt-tRNA<sup>Tyr</sup> and the human synthetase. The NMR-structure of the isolated domain from *B. stearrowthermophilus* (Guijarro et al., 2002), the crystal structure of the *T. thermophilus* TyrRS in complex with tRNA<sup>Tyr</sup> (Yaremchuk et al., 2002), and the property of the N-terminal domain of eukaryal class IIb synthetases, e.g. yeast AspRS, to reinforce tRNA binding (Frugier et al., 2000), bring support to this view.

The crystals used here could only be grown with synthetase purified under reducing conditions. Likewise, activity of human mt-TyrRS- $\Delta$ S4 (as well as full-length TyrRS) requires reducing conditions, indicating the necessity of non-oxidized cysteine residues for crystallization and functionality. Cysteine residues were proposed to participate in the catalytic mechanism of yeast TyrRS, based on the observation that modification of two reactive sulfhydryl groups per subunit leads to inactivation of the enzyme (Faulhammer and Cramer, 1977). Given the location of the cysteine residues in the structure of human mt-TyrRS (Figure 2), the role of these residues cannot be directly correlated with the catalytic mechanism of tyrosylation. Rather, these cysteines affect activity indirectly by promoting formation of an inactive synthetase conformation. Note that the mitochondrial genetic machinery, including mt-TyrRS, is located in the matrix, where conditions are overall reducing (Korge and Weiss, 2006).

Aminoacyl-tRNA synthetases are characterized by an intrinsic structural plasticity leading to high mobility of certain structural domains (reviewed in Ibba et al., 2005). This propensity of conformational flexibility, often linked with enzyme function, is well documented in TyrRS

structures (reviewed by Bedouelle, 2005) and finds support with the properties of human mt-TyrRS. In particular the "KMSKS"-loops are either disordered and not visible in electron density or present high B-factors as compared to the whole molecule. This is due to high mobility during each steps of the tyrosylation reaction. In the adenylate-bound form of human mt-TyrRS-DS4, the "KMSKS"-loop is rather remote from the active site, thus explaining a relative lack of constraints in the structure. Further, many functional studies concluded for a functional asymmetry of homodimeric TyrRSs (Bosshard et al., 1975; Ward and Fersht, 1988) and their half-of-the-sites activity is well established (Bedouelle, 2005). However, most TyrRS structures show crystallographic symmetry and ligand binding to both subunits. This discrepancy was discussed previously and explained by the duration of the crystallization assays that allows the TyrRS binding sites to be filled by their ligands (Yaremchuk et al., 2002). With the human mt-TyrRS, we clearly observe a structural asymmetry of the enzyme in complex with Tyr-AMS. It concerns essentially two regions in the catalytic domain and one region in the anticodon-binding domain (Figure 4). Although part of this asymmetry could be due to packing effects, the present structure clearly shows an intrinsic plasticity of the enzyme in complex with the adenylate analog. We conjecture that structural perturbations propagate from the dimer interface towards the two catalytic sites and the a-ACB modules at the protein distal extremities. A similar behavior was recently reported for the 322-residue long C-truncated *E. coli* TyrRS, but restricted to the sole structural asymmetry in the "KMSKS"-loops (Kobayashi et al., 2005).

So far, no suitable crystals of the mt-TyrRS•tRNA complex could be grown (Bonfond et al., 2007) and only partial answers can be given about the tRNA binding over the two subunits, like in other TyrRSs (Figure 4) (Bedouelle, 2005). Docking of tRNA (in the conformation seen in *T. thermophilus* complex, Yaremchuk et al., 2002) on the human enzyme necessitates structural distortions on the two macromolecules to allow both extremities of the tRNA to contact simultaneously the catalytic and anticodon binding domains. This is in line with the structural plasticity of mammalian mt-tRNAs allowing the two branches of the L to vary their angle (Helm et al., 2000). Despite important sequence variability in the a-ACB domain, we notice the presence in the human enzyme of residue Asp311, homologous to Asp259 in *T. thermophilus* TyrRS that interacts with the anticodon base G34 of cognate tRNA<sup>Tyr</sup> (Yaremchuk et al., 2002). Other amino acids identified as anticodon binders are not conserved in human mt-TyrRS. As to the binding site of the acceptor arm, remember that human mt-TyrRS is functionally characterized by its lack of specificity towards base-pair N1–N72 in tRNA<sup>Tyr</sup>, in contrast to eubacterial, archaeal and eukaryal TyrRSs that recognize either G1–C72 or C1–G72 pairs, implying that the N1–N72



pair does not belong to the tyrosine identity set in human mitochondria. In other words only A73 and the anticodon triplet GUA specify tyrosylation (Bonnefond et al., 2005c). On the other hand, previous sequence and structure analysis of TyrRSs revealed a mosaic nature of the two clusters whose amino acids are in vicinity or in contact with the tRNA<sup>Tyr</sup> acceptor end (Bonnefond et al., 2005b). To shed light towards understanding tRNA<sup>Tyr</sup> binding on the catalytic domain, we undertook a mutagenesis analysis. Our analysis is completed by former studies on the phylogenetically related *B. stearothermophilus* TyrRS (Labouze and Bedouelle, 1989). Notably, both studies pinpoint the functional importance of an amino acid from the binding site, namely Ser200, a conserved mitochondrial amino acid in the human enzyme and its homolog Ala150 in the *Bacillus* TyrRS. Altogether, the present knowledge on the three-dimensional structure and functionality of human mt-TyrRS are in light with the sequence-based phylogeny of TyrRSs (Bonnefond et al., 2005c) and the evolutionary relationship between mitochondria and eubacteria. More precisely, the relaxed tyrosine identity in human mitochondria, a property likely idiosyncratic to vertebrate mt-TyrRSs, would be achieved by a decreased number of specific contacts between tRNA and synthetase and not by the presence of archaeal features (i.e. Met252) as suggested by sequence comparisons (Bonnefond et al., 2005b). Additional mutagenesis and crystallographic studies on the complex with tRNA<sup>Tyr</sup> are required to fully understand these functional peculiarities. Interestingly, altered identity elements were already described in other mitochondrial systems (Lovato et al., 2001; Chimnarok et al., 2005). It remains to be investigated whether the strategy of reducing the number of identity elements is employed by other mitochondrial synthetases. Structure-function studies on human mt-AspRS support this interpretation (Fender et al., 2006).

## **EXPERIMENTAL PROCEDURES**

### **Cloning, Expression, Mutagenesis and Purification of mt-TyrRSs**

Top10 *E. coli* strains containing plasmids of full-length or truncated TyrRS were grown in LB medium and TyrRS expression induced with IPTG. Recovered cells, suspended in a phosphate buffer at pH 8.0, were sonicated and resulting supernatant resolved with an imidazole gradient on a Ni-NTA column (Qiagen). Enzyme fractions were dialyzed against HEPES-NaOH pH 6.7, NaCl 300 mM, DTE 10 mM and 50% glycerol. Glycerol was removed from the concentrated protein solution by two successive dialyses against the same buffer but with 10% glycerol and finally without glycerol. Experimental details and exact sequence

at both extremities of the cloned enzyme were described previously (Bonfond et al., 2005a; Bonfond et al., 2007).

The three variants of full-length human mt-TyrRS (Ser200Glu, Gln202Ala and Met252Ala) were prepared using QuickChange® site-directed mutagenesis kit (Stratagene). Protein expression and purification were conducted under reducing conditions as indicated above.

### Functional Assays

ATP/PPi exchange was done in media containing 100 mM HEPES-NaOH pH 7.2, 10 mM MgCl<sub>2</sub>, 2 mM KF, 2 mM ATP and 2 mM [<sup>32</sup>P]PPi (1–2 cpm/pmol). Reactions were initiated by adding 1 mg of the appropriate mt-TyrRS version. The [<sup>32</sup>P]ATP formed after incubation at 37°C was determined as described (Campanacci et al., 2004). Controls with no TyrRS or no amino acid were conducted in parallel.

Aminoacylation of native tRNA<sup>Tyr</sup> or mt-tRNA<sup>Tyr</sup> transcripts was performed as described (Bonfond et al., 2005b). Kinetic parameters  $K_M$  and  $k_{cat}$  were determined from Lineweaver-Burk plots. Displayed data are averages of at least 2 independent experiments.

### Crystallization and Structure Determination

Initial tetragonal crystals were obtained by vapor-diffusion (sitting drops of 600 nL to 2 mL) in CrystalQuick™ microplates (Greiner Bio-One) prefilled with EasyXtal screens (Qiagen) (Bonfond et al., 2007). Starting from these conditions 30% PEG 4000 w/v, 200 mM NH<sub>4</sub> acetate, 100 mM Na acetate pH 4.6), a new orthorhombic crystal form with a prismatic habit was obtained in the presence of 5 mM Tyr-AMS (RNA-tec, Belgium) by rising the pH from 4.6 to 6.5 by addition of 100 mM Tris-HCl pH 7.5 (from Optisalt™ additive screen, Qiagen). These crystals of 200 μm in their largest dimension diffract X-rays to 2.2 Å resolution.

Crystals were mounted in cryoloops (Hampton Research) and flash-frozen in a nitrogen stream after brief soaking in cryoprotecting Paratone® (Hampton Research). They were analyzed at 100°K on beamline ID23-1 (ESRF, France) equipped with an ADSC Quantum 315 CCD detector. Data were processed with the XDS package (Kabsch, 1993). Analysis of solvent content performed with the CCP4 package (Collaborative Computational Project, 1994)(Collaborative Computational Project, 1994) gave a unique solution consisting of one or two polypeptide chain(s) per asymmetric unit in the case of native (tetragonal) or Tyr-AMS substituted (orthorhombic) crystals, respectively. A molecular-replacement (MR) solution was found for the original tetragonal crystals (space group  $P4_32_12$ ) using the structure of the closely related *E. coli* ortholog (Bonfond et al., 2007) but the last 80 C-terminal residues (corresponding to the α-ACB module) could not be built due to the lack of interpretable

electron density. Finally, the MR solution after partial refinement was used as a starting point for MR using orthorhombic data. In contrast to the tetragonal situation, a near to complete chain could be traced for the two monomers of mt-TyrRS- $\Delta$ S4 in the orthorhombic unit cell at a resolution of 2.2 Å. Model building and refinement were carried out using Coot (Emsley and Cowtan, 2004) and CNS (Brünger et al., 1998), respectively. The final structure includes 317+324 out of 2x356 residues [lacking regions 28–36, 66–72, 276–288 and 374–383 (including the 6-His tag) in monomer A and 28–36, 276–288 and 374–383 (with the 6-His tag) in monomer B], two adenylate analogs and 90 water molecules. Data collection and refinement statistics are given in Table 1. R.m.s.d. were calculated with PyMol and LSQMAN (DeLano, 2002; Kleywegt and Jones, 1994). Electrostatic potential surfaces were represented using APBS module (Baker et al., 2001) of PyMol. Docking of tRNA on TyrRS was achieved using the SSM algorithm (Krissinel and Henrick, 2004) and the tRNA structure as seen in the complex with *T. thermophilus* TyrRS (Yaremchuk et al., 2002).

### **Supplemental Data**

Supplemental Data, including two figures and sequence alignment of mitochondrial TyrRSs, are available online at <http://www.structure.org>.

### **ACKNOWLEDGEMENTS**

We thank M. Sissler for stimulating discussions, G. Bec for help in the use of robotics, as well as P. Dumas, E. Ennifar and V. Oliéric for advices during the crystallographic steps of this work. We thank C. Paulus for excellent technical help for mutagenesis and A. Fender for discussions at the early stages of this work. This study was supported by the Centre National de la Recherche Scientifique (CNRS), Université Louis Pasteur (Strasbourg) and the French Ministry for Research (ACI "BCMS" 042358). L.B. and E.T. were supported by doctoral grants from the French Ministry for Research and C.S. was the recipient of a Marie Curie European Reintegration Grant (MERC-CT-2004-004898).

**Received:** 20 April 2007

## REFERENCES

- Abergel, C., Chenivesse, S., Byrne, D., Suhre, K., Arondel, V. and Claverie, J.-M. (2005) Mimivirus TyrRS: preliminary structural and functional characterization of the first aminoacyl-tRNA synthetase found in a virus. *Acta Crystallogr. F* **61**, 212–215.
- Akins, R.A. and Lambowitz, A.M. (1987). A protein required for splicing group I introns in *Neurospora* mitochondria is mitochondrial tyrosyl-tRNA synthetase or a derivative thereof. *Cell* **50**, 331–345.
- Antonicka, H., Sasarman, F., Kennaway, N.G., Shoubridge, E.A. (2006). The molecular basis for tissue specificity of the oxidative phosphorylation deficiencies in patients with mutations in the mitochondrial translation factor EFG1. *Hum. Mol. Genet.* **15**, 1835–1846.
- Baker, N.A., Sept, D., Joseph, S., Holst, M.J. and McCammon, J.A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10037–10041.
- Bedouelle, H. and Winter, G. (1986). A model of synthetase/transfer RNA interaction as deduced by protein engineering. *Nature* **320**, 371–373.
- Bedouelle, H. (2005). Tyrosyl-tRNA Synthetases. In *Aminoacyl-tRNA synthetases*, M. Ibba, C. Francklyn and S. Cusack, eds. (Landes Biosciences, Georgetown, TX), pp. 111–124.
- Beuning, P.J. and Musier-Forsyth, K. (1999). Transfer RNA recognition by aminoacyl-tRNA synthetases. *Biopolymers* **52**, 1–28.
- Bonnefond, L., Fender, A., Rudinger-Thirion, J., Giegé, R., Florentz, C. and Sissler, M. (2005a). Towards the full set of human mitochondrial aminoacyl-tRNA synthetases: characterization of AspRS and TyrRS. *Biochemistry* **44**, 4805–4816.
- Bonnefond, L., Frugier, M., Giegé, R. and Rudinger-Thirion, J. (2005b). Human mitochondrial TyrRS disobeys the tyrosine identity rules. *RNA* **11**, 558–562.
- Bonnefond, L., Giegé, R. and Rudinger-Thirion, J. (2005c). Evolution of the tRNA<sup>Tyr</sup>/TyrRS aminoacylation systems. *Biochimie* **87**, 873–883.
- Bonnefond, L., Frugier, M., Touzé, E., Lorber, B., Florentz, C., Giegé, R., Rudinger-Thirion, J. and Sauter, C. (2007). Tyrosyl-tRNA synthetase: the first crystallization of a human mitochondrial aminoacyl-tRNA synthetase. *Acta Crystallogr. F* **63**, 338–341.
- Bosshard, H.R., Koch, G.L.E. and Hartley, B.S. (1975). Aminoacyl-tRNA synthetases from *Bacillus stearothermophilus*. Asymmetry of substrate binding to tyrosyl-tRNA synthetase. *Eur. J. Biochem.* **53**, 493–498.
- Brandon, M., Baldi, P. and Wallace, D.C. (2006). Mitochondrial mutations in cancer. *Oncogene* **25**, 4647–4662.
- Brick, P., Bhat, T.N. and Blow, D.M. (1989). Structure of tyrosyl-tRNA synthetase refined at 2.3 Å resolution. Interaction of the enzyme with the tyrosyl adenylate intermediate. *J. Mol. Biol.* **208**, 83–98.
- Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. (1998). Crystallography and NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921.
- Campanacci, V., Dubois, D.Y., Becker, H.D., Kern, D., Spinelli, S., Valencia, C., Pagot, F., Salomoni, A., Grisel, S., Vincentelli, R., et al. (2004). The *Escherichia coli* YadB gene product reveals a novel aminoacyl-tRNA synthetase like activity. *J. Mol. Biol.* **337**, 273–283.

- Cherniack, A.D., Garriga, G., Kittle Jr., J.D., Akins, R.A. and Lambowitz, A.M. (1990). Function of *Neurospora* mitochondrial tyrosyl-tRNA synthetase in RNA splicing requires an idiosyncratic domain not found in other synthetase. *Cell* **62**, 745–755.
- Chimnaronk, S., Gravers Jeppesen, M., Shimada, N., Suzuki, T., Nyborg, J. and Watanabe, K. (2004). Crystallization and preliminary X-ray diffraction study of mammalian mitochondrial seryl-tRNA synthetase. *Acta Crystallogr. D* **60**, 1319–1322.
- Chimnaronk, S., Gravers Jeppesen, M., Suzuki, T., Nyborg, J. and Watanabe, K. (2005). Dual-mode recognition of noncanonical tRNAs(Ser) by seryl-tRNA synthetase in mammalian mitochondria. *EMBO J.* **24**, 3369–3379.
- Collaborative Computational Project, N., 1994. (1994). The CCP4 Suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763.
- DeLano, W.L. (2002). The PyMOL molecular graphics system. Delano Scientific (<http://www.pymol.org>), San Carlos, CA, USA.
- Emsley, P. and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132.
- Faulhammer, H.G. and Cramer, F. (1977). Tyrosyl-tRNA synthetase from baker's yeast. Rapid isolation by affinity elution, molecular weight of the enzyme, and determination of essential sulfhydryl groups. *Eur. J. Biochem.* **75**, 561–570.
- Fechter, P., Rudinger-Thirion, J., Théobald-Dietrich, A. and Giegé, R. (2000). Identity of tRNA for yeast tyrosyl-tRNA synthetases: tyrosylation is more sensitive to identity nucleotides than to structural features. *Biochemistry* **39**, 1725–1733
- Fechter, P., Rudinger-Thirion, J., Théobald-Dietrich, A. and Giegé, R. (2001). Major tyrosine identity determinants in *Methanococcus jannaschii* and *Saccharomyces cerevisiae* tRNA<sup>Tyr</sup> are conserved but expressed differently. *Eur. J. Biochem.* **268**, 761–767.
- Fender, A., Sauter, C., Messmer, M., Pütz, J., Giegé, R., Florentz, C. and Sissler, M. (2006). Loss of a primordial identity element for a mammalian mitochondrial aminoacylation system. *J. Biol. Chem.* **281**, 15980–15986.
- Florentz, C., Ohm, B., Tryoen-Tóth, P., Pütz, J. and Sissler, M. (2003). Human mitochondrial tRNAs in health and disease. *Cell. Mol. Life Sci.* **60**, 1356–1375.
- Frugier, M., Moulinier, L and Giegé, R. (2000). A domain in the N-terminal extension of class IIb eukaryotic aminoacyl-tRNA synthetases important for tRNA binding. *EMBO J.* **19**, 2371–2380
- Giegé, R., Sissler, M. and Florentz, C. (1998). Universal rules and idiosyncratic features in tRNA identity. *Nucleic Acids Res.* **26**, 5017–5035.
- Guijarro, J.I., Pintar, A., Prochnicka-Chalufour, A., Guez, V., Gilquin, B., Bedouelle, H. and Delepierre M. (2002). Structure and dynamics of the anticodon arm binding domain of *Bacillus stearothermophilus* tyrosyl-tRNA synthetase. *Structure* **10**, 31–317.
- Helm, M., Brulé, H., Friede, D., Giegé, R., Pütz, J. and Florentz, C. (2000). Search for characteristic features of mammalian mitochondrial tRNAs. *RNA* **6**, 1356–1379.
- Ibba, M., Francklyn, C. and Cusack, S. (eds) (2005). Aminoacyl-tRNA synthetases (Landes Biosciences, Georgetown, TX).
- Irwin, M.J., Nyborg, J., Reid, B.R. and Blow, D.M. (1976). The crystal structure of tyrosyl-transfer synthetase at 2.7 Å resolution. *J. Mol. Biol.* **105**, 577–586.
- Jacobs, H.T. and Turnbull, D.M. (2005). Nuclear genes and mitochondrial translation: a new class of genetic disease. *Trends Genet.* **21**, 312–314.

- Kabsch, W. (1993). Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Cryst.* **26**, 795–800.
- Kleeman, T.A., Wei, D., Simpson, K.L. and First, E.A. (1997). Human tyrosyl-tRNA synthetase shares amino acid sequence homology with a putative cytokine. *J. Biol. Chem.* **272**, 14420–14425.
- Kleywegt, G.J. and Jones, T.A. (1994). A super position. *Newsletter Protein Crystallogr.* **31**, 9–14.
- Kobayashi, T., Nureki, O., Ishitani, R., Yaremchuk, A., Tukalo, M., Cusack, S., Sakamoto, K. and Yokoyama, S. (2003). Structural basis for orthogonal tRNA specificities of tyrosyl-tRNA synthetases for genetic code expansion. *Nat. Struct. Biol.* **10**, 425–432.
- Kobayashi, T., Takimura, T., Sekine, R., Kelly, V.P., Kamata, K., Sakamoto, K., Nishimura, S. and Yokoyama, S. (2005). Structural snapshots of the KMSKS loop rearrangement for amino acid activation by bacterial tyrosyl-tRNA synthetase. *J. Mol. Biol.* **346**, 105–117.
- Korge, P. and Weiss, J.N. (2006). Redox regulation of endogeneous substrate oxidation by cardiac mitochondria. *Am. J. Physiol. Heart Circ. Physiol.* **291**, H1436–H1445.
- Krissinel, E. and Henrick, K. (2004). Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr. D* **60**, 2256–2268.
- Kuratani, M., Sakai, H., Takahashi, M., Yanagisawa, T., Kobayashi, T., Murayama, K., Chen, L., Liu, Z.J., Wang, B.C., Kuroishi, C., et al. (2006). Crystal structures of tyrosyl-tRNA synthetases from *Archaea*. *J. Mol. Biol.* **355**, 395–408.
- Labouze, E. and Bedouelle, H. (1989). Structural and kinetic bases for the recognition of tRNA<sup>Tyr</sup> by tyrosyl-tRNA synthetase. *J. Mol. Biol.* **205**, 729–735.
- Laskowsky, R.A., Rullmann, J.A., MacArthur M.W., Kaptein, R. and Thornton, J.M. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* **8**, 477–486.
- Lee, C.P. and RajBhandary, U.L. (1991). Mutants of *Escherichia coli* initiator tRNA which suppress amber codons in *Saccharomyces cerevisiae* and are aminoacylated with tyrosine by yeast extracts. *Proc. Natl. Acad. Sci. U.S.A.* **88**, 11378–11382.
- Lovato, M.A., Chihade, J.W. and Schimmel, P. (2001). Translocation within the acceptor helix of a major tRNA identity determinant. *EMBO J.* **20**, 4848–4853.
- Nair, S., Ribas de Pouplana, L., Houtman, F., Avruch, A., Shen, X. and Schimmel, P. (1997). Species-specific tRNA recognition in relation to tRNA synthetase contact residues. *J. Mol. Biol.* **269**, 1–9.
- Paukstelis, P.J., Coon, R., Madabusi, L., Nowakowski, J., Monzingo, A., Robertus, J. and Lambowitz, A.M. (2005). A tyrosyl-tRNA synthetase adapted to function in group I intron splicing by acquiring a new RNA binding surface. *Mol. Cell* **17**, 417–428.
- Qiu, X., Janson, C.A., Smith, W.W., Green, S.M., McDevitt, P., Johanson, K., Carter, P., Hibbs, M., Lewis, C., Chalker, A., et al. (2001). Crystal structure of *Staphylococcus aureus* tyrosyl-tRNA synthetase in complex with a class of potent and specific inhibitors. *Protein Sci.* **10**, 2008–2016.
- Quinn, C.L., Tao, N. and Schimmel, P. (1995). Species-specific microhelix aminoacylation by a eukaryotic pathogen tRNA synthetase dependent on a single base pair. *Biochemistry* **34**, 12489–12495.
- Scheper, G.C., van de Kloek, T., van Andel, R., van Berkel, C.G.M., Sissler, M., Smet, J., Muravina, T.I., Serkov, S.V., Uziel, G., Bugiani, M., et al. (2007). Mitochondrial aspartyl-tRNA

synthetase deficiency causes leukoencephalopathy with brain stem and spinal cord involvement and lactate elevation. *Nature Genet.* **39**, 534–539.

Sprinzi, M. and Vassilenko, K.S. (2005). Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res.* **33**, D139–D140.

Taylor, R.W. and Turnbull, D.M. (2005). Mitochondrial DNA mutations in human disease. *Nat. Rev. Genet.* **6**, 389–402.

Wakasugi, K., Quinn, C.L., Tao, N. and Schimmel, P. (1998). Genetic code in evolution: switching species-specific aminoacylation with a peptide transplant. *EMBO J.* **17**, 297–305.

Wakasugi, K. and Schimmel, P. (1999). Two distinct cytokines released from a human aminoacyl-tRNA synthetase. *Science* **284**, 147–150.

Ward, W.H. and Fersht, A.R. (1988). Tyrosyl-tRNA synthetase acts as an asymmetric dimer in charging tRNA. A rationale for half-of-the-sites activity. *Biochemistry* **27**, 5525–5530.

Wolf, Y.I., Aravind, L., Grishin, N.V. and Koonin, E.V. (1999). Evolution of aminoacyl-tRNA synthetases--analysis of unique domain architectures and phylogenetic trees reveals a complex history of horizontal gene transfer events. *Genome Res.* **9**, 689–710.

Xin, Y., Li, W. and First, E.A. (2000). The 'KMSKS' motif in tyrosyl-tRNA synthetase participates in the initial binding of tRNA(Tyr). *Biochemistry* **39**, 340–347.

Yang, X.L., Skene, R.J., McRee, D.E. and Schimmel, P. (2002). Crystal structure of a human aminoacyl-tRNA synthetase cytokine. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 15369–15374.

Yang, X.L., Liu, J., Skene, R.J., McRee, D.E. and Schimmel, P. (2003). Crystal structure of an EMAP-II-like cytokine released from a human tRNA synthetase. *Helvetica Chimica Acta* **86**, 1246–1257.

Yaremchuk, A., Krikliivi, I., Tukalo, M. and Cusack, S. (2002). Class I tyrosyl-tRNA synthetase has a class II mode of cognate tRNA recognition. *EMBO J.* **21**, 3829–3840.

Zhang, Y., Wang, L., Schultz, P.G. and Wilson, I.A. (2005). Crystal structures of apo wild-type *M. jannaschii* tyrosyl-tRNA synthetase (TyrRS) and an engineered TyrRS specific for O-methyl-L-tyrosine. *Protein Sci.* **14**, 1340–1349.

## Accession Numbers

The atomic coordinates and structure factors of human mitochondrial TyrRS- $\Delta$ S4 in complex with an adenylate analog have been deposited in the Protein Data Bank (PDB ID code 2PID).

**Table 1.** Crystal analysis and structure refinement of human mt-TyrRS- $\Delta$ S4 in complex with two Tyr-AMS ligands.

<b>Crystal analysis</b>	
Synchrotron beamline	ID23-1 (ESRF)
Wavelength (Å)	0.97
Space group	$P2_12_12_1$
Unit cell a, b, c parameters (Å)	54.0, 62.6, 195
Crystal mosaicity (deg)	0.64
Resolution range (Å)	2.2–20 (2.2–2.3)
Number of observations	1219567
Number of unique reflections	32590
Completeness (%)	99.6 (100)
Multiplicity	11 (10)
$R_{\text{merge}}$ (%)	11.2 (50)
$I/\sigma(I)$	22 (10.5)
Solvent content (%)	40
Asymmetric unit content	2 monomers
<b>Structure refinement</b>	
R-factor (%)	19.6 (22.8)
free R-factors (%)	24.4 (30.9)
Number of atoms	
protein	5104
ligand	70
solvent	90
B-factors (Å <sup>2</sup> )	
overall	40.4
protein	40.8
ligand	32.5
solvent	39.9
R.m.s.d. for bond distances (Å), angles (°)	0.008, 1.4
Ramachandran plot (%)*	
most favored region	94.7
additionally allowed region	4.9
generously allowed region	0.4

(\*) Statistics from PROCHECK (Laskowski et al., 1996)

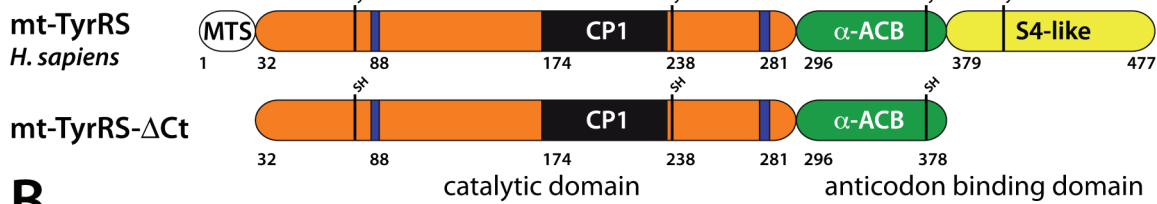
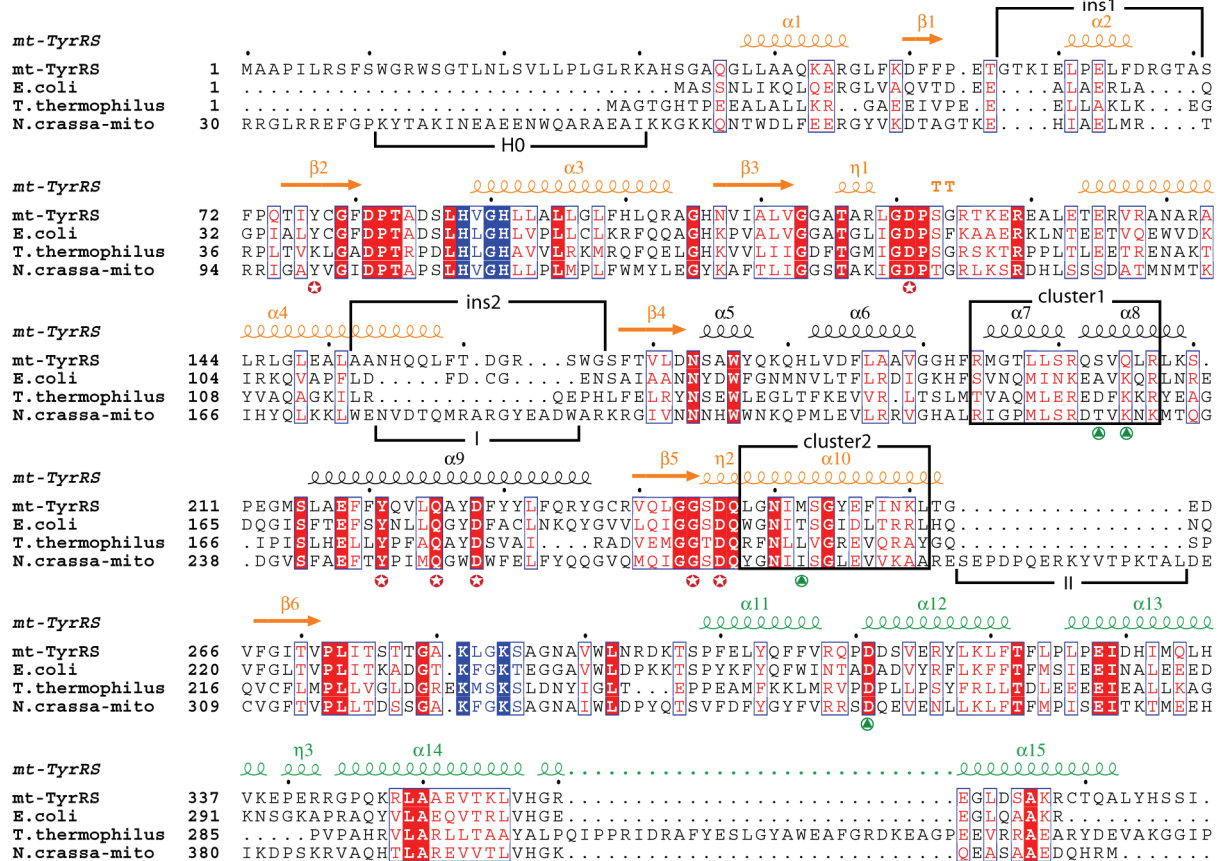


**Table 2.** Kinetic parameters of the tyrosylation reaction of human mitochondrial tRNA<sup>Tyr</sup> transcript by native and mutated human mt-TyrRS

mt-TyrRS	mt-tRNA <sup>Tyr</sup>			
	$K_M$ ( $\mu\text{M}$ )	$k_{\text{cat}}$ ( $10^{-3} \cdot \text{s}^{-1}$ )	$k_{\text{cat}}/K_M$ ( $10^{-3} \cdot \text{s}^{-1} \cdot \mu\text{M}^{-1}$ )	L (x-fold)
native	1.5	46	31	1
Ser <b>200</b> Glu	nm	nm	nm	nm
Gln <b>202</b> Ala	2.7	140	52	1.7
Met <b>252</b> Ala	1.7	45	26	1.2

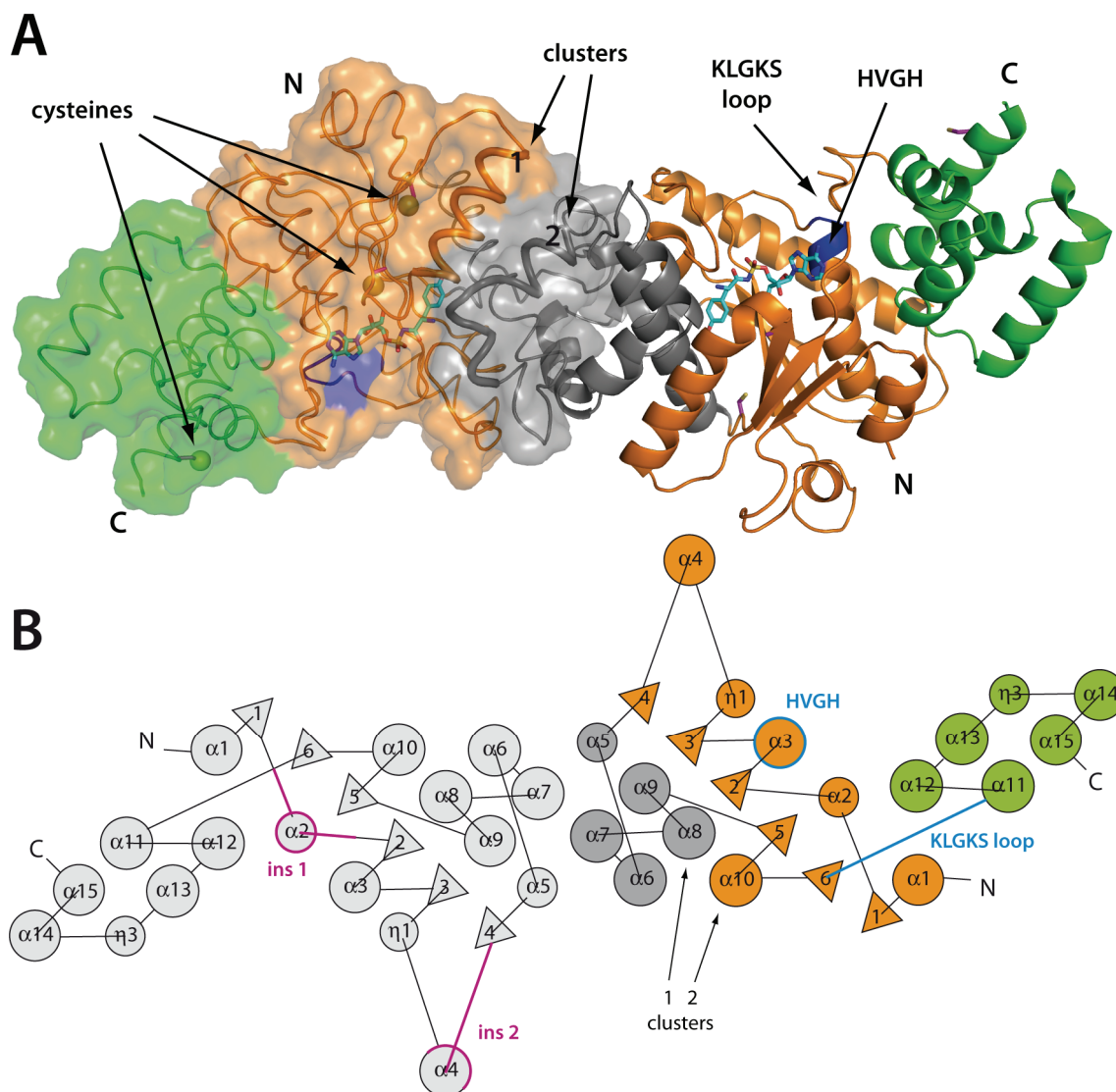
All proteins are fully active in ATP/PPi exchange assay; nm: not measurable.

L = loss in tyrosylation efficiency.

**A****B****Figure 1. Sequence Features of Human Mitochondrial TyrRS**

(A) Modular organization of human mt-TyrRS and mt-TyrRS-ΔS4. The catalytic domain (orange), with the location of the HVGH and KLGKS signature sequences (blue), is disrupted by the CP1 (Connective Peptide 1) in black. The anticodon binding region includes the helical α-ACB (AntiCodon Binding) domain (green) and the S4-like (S4 ribosomal protein-like) domain (yellow) for entire mt-TyrRS. The Nt-located mitochondrial targeting sequence (MTS) in the native enzyme is shown. Location of cysteine residues is indicated by SH-labeled lines.

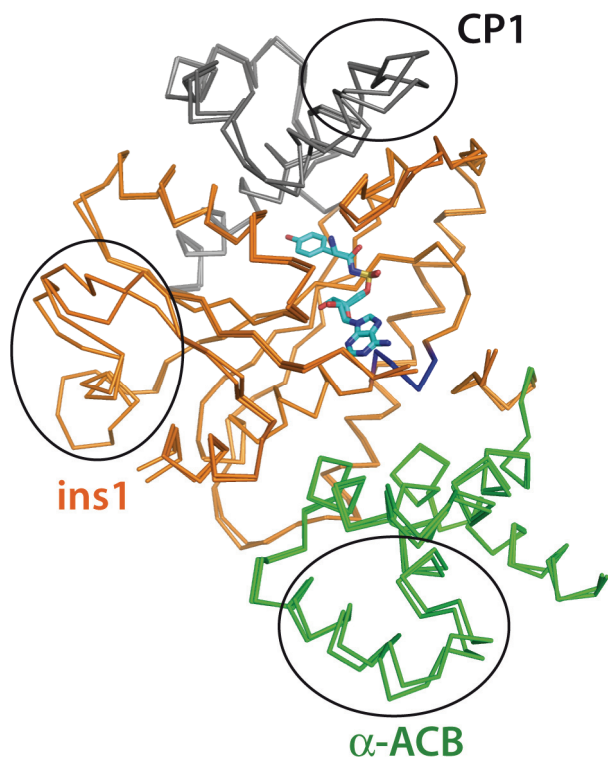
(B) Structure-based alignment of human mt-TyrRS sequence (including MTS domain) with TyrRSs from *E. coli* (PDB code 1VBM), *T. thermophilus* (PDB code 1H3E) and *N. crassa* mitochondria (PDB code 1Y42). Conserved and similar residues are highlighted in red. The secondary structure elements as found in the crystal structure of mt-TyrRS-ΔS4 (see Figure 2) are shown above its sequence. Conserved and semi-conserved residues in class I synthetase canonical signature "HIGH" and "KMSKS" sequences are highlighted in blue (blue background and blue scripts, respectively). Sequences of cluster 1 (residues 194–207) and cluster 2 (residues 248–261) are framed in the alignment. Specific insertions in human (ins1, ins2) and *N. crassa* (H0, I, II) mt-TyrRSs are indicated above and under the alignment, respectively. Strategic amino acids discussed in the text are pointed (in red for small substrates, see Figure 6C, and in green for tRNA substrate, see Figure 7A).



**Figure 2. Crystallographic Structure of Human mt-TyrRS- $\Delta$ S4**

(A) Overall crystallographic structure of dimeric mt-TyrRS- $\Delta$ S4 showing the elongated shape of the protein (color code as in Figure 1A) in its Tyr-AMS bound form. One monomer is drawn in ribbon with a transparent surface and clusters 1 and 2 as well as the three cysteine residues emphasized. The other monomer is drawn as a cartoon with the location of the KLGKS-loop indicated close to the HVGH-motif (in blue) on top of helix  $\alpha$ 3.

(B) Topology diagram of the dimeric synthetase with the secondary structure elements labeled ( $\alpha$ -helices shown as circles, and  $\beta$ -strands as triangles) and one monomer (at the right) colored as in Figure 1. The diagram specifies the location of the signature sequences HVGH on  $\alpha$ 3 and KLGKS in the loop joining  $\beta$ 6 and  $\alpha$ 11 (in blue on monomer B at the right), as well as of "ins1" between  $\beta$ 1 and  $\beta$ 2 and "ins2" between  $\alpha$ 4 and  $\beta$ 4 (in red on monomer A at the left).

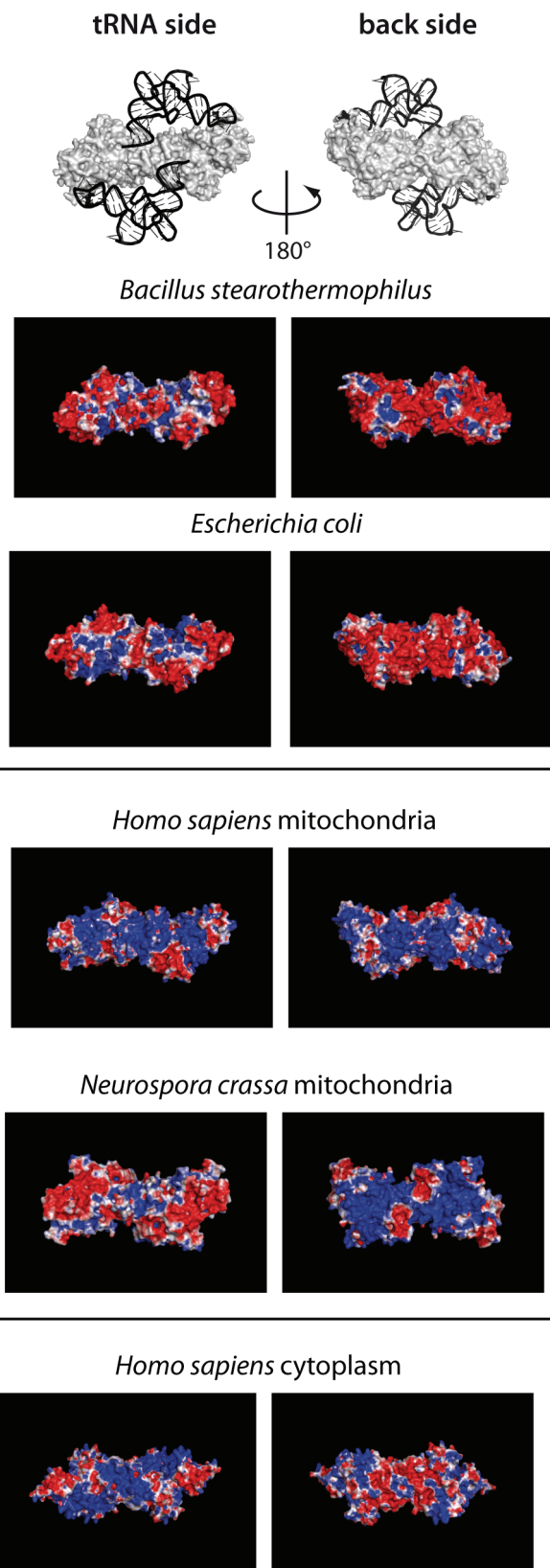


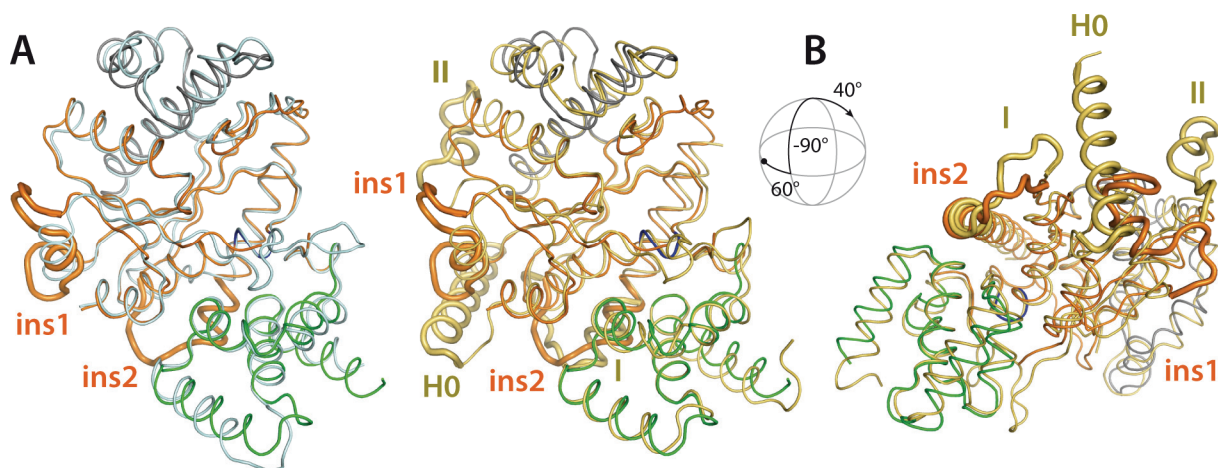
### Figure 3. Conformational Asymmetry in Homodimeric Human mt-TyrRS- $\Delta$ S4

Superimposition of the two subunits of mt-TyrRS- $\Delta$ S4 in complex with Tyr-AMS (backbones colored as in Figure 1A, in heavy and light colors for monomer A and B, respectively). The regions with largest asymmetries are circled. Note that the two adenylate analogs (in blue) almost perfectly superimpose.

### Figure 4. Electrostatic Surface Potentials of Several mt-TyrRS- $\Delta$ S4

Blue, white and red regions correspond to positive, neutral and negative potentials, respectively. Computations used the following coordinates: PDB codes 3TS1, 1WQ3, 2PID (this work), 1Y42, 1N3L. Two orientations of the TyrRSs (all Ct-truncated) are displayed: the side where the two tRNA molecules bind to the synthetase (on the left) and the back orientation rotated by 180° (on the right). The two mitochondrial enzymes are framed. Location of the two tRNA molecules bridging the TyrRS subunits is indicated, as found in the crystal structure from *T. thermophilus* (Yaremchuk et al., 2002).

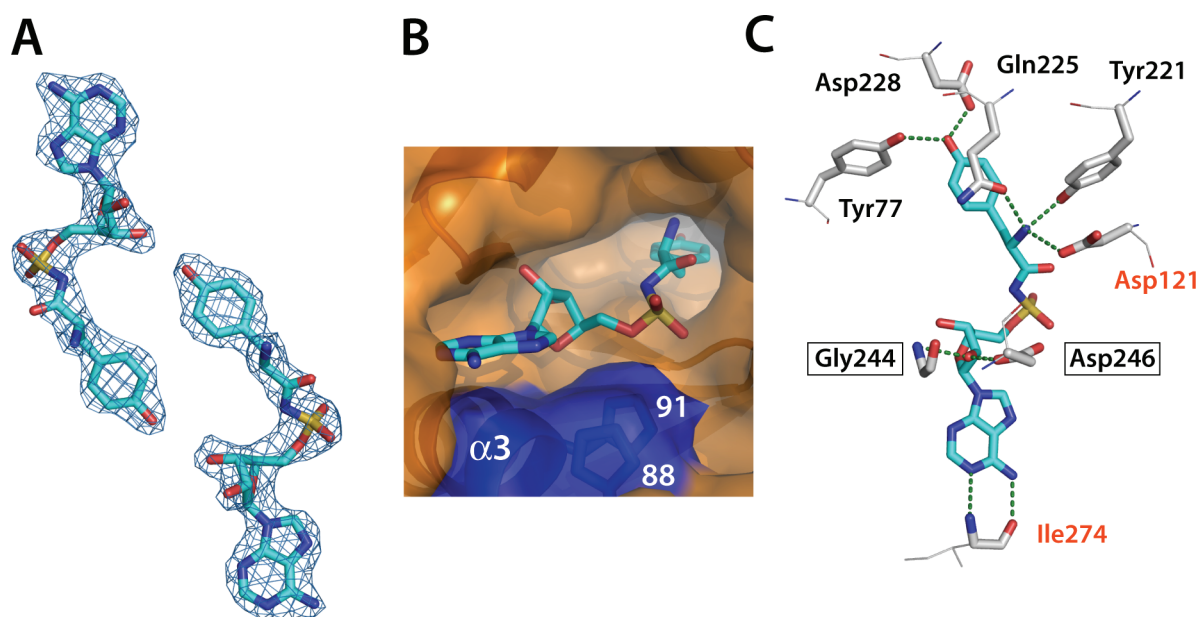




### Figure 5. Structural Idiosyncrasies in Human mt-TyrRS- $\Delta$ S4

(A) Comparison of the TyrRS- $\Delta$ S4 subunits of human mitochondria (monomer B, this work) and *E. coli* (PDB code 1VBM). Backbone of mt-TyrRS is colored as in Figure 1A and that of *E. coli* TyrRS in grey.

(B) Comparison of the mt-TyrRS- $\Delta$ S4 subunits of human (monomer B, this work) and *N. crassa* (PDB code 1Y42). Human mt-TyrRS is colored as in panel A and that from *N. crassa* in yellow. To better emphasize the structural differences, the superimpositions are displayed under two orientations. Notice that *N. crassa*-specific elements II and H0 (colored in yellow) are absent in the human protein.



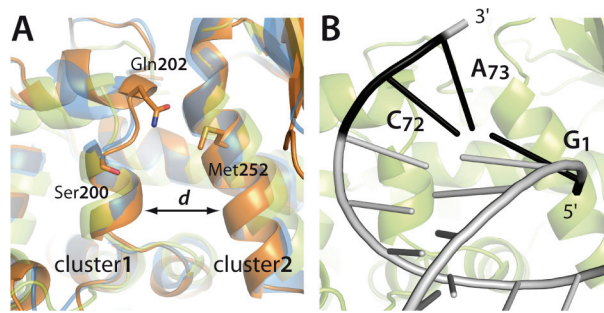
### Figure 6. Snapshot in the Active Site of Dimeric Human mt-TyrRS- $\Delta$ S4

(A) Electron density of the two Tyr-AMS adenylate analogs in complex with mt-TyrRS- $\Delta$ S4 (from monomer A on the left and monomer B on the right).

(B) Active site pocket with the bound Tyr-AMS and the HVGH domain on top of helix- $\alpha$ 3 emphasized in blue. Notice, the proximity of the adenine ring from Tyr-AMS with the imidazole ring from His88 and His91, the two residues from the HVGH signature sequence.

(C) Cartoon of the well resolved adenylate analog (in blue) and the network of nine H-bond contacts it makes with the eight essential amino acids from the catalytic site. Amino acids specific to class I synthetases are boxed; those strictly conserved in TyrRSs are labeled in bold and those specific to mitochondrial and eubacterial TyrRSs in red.

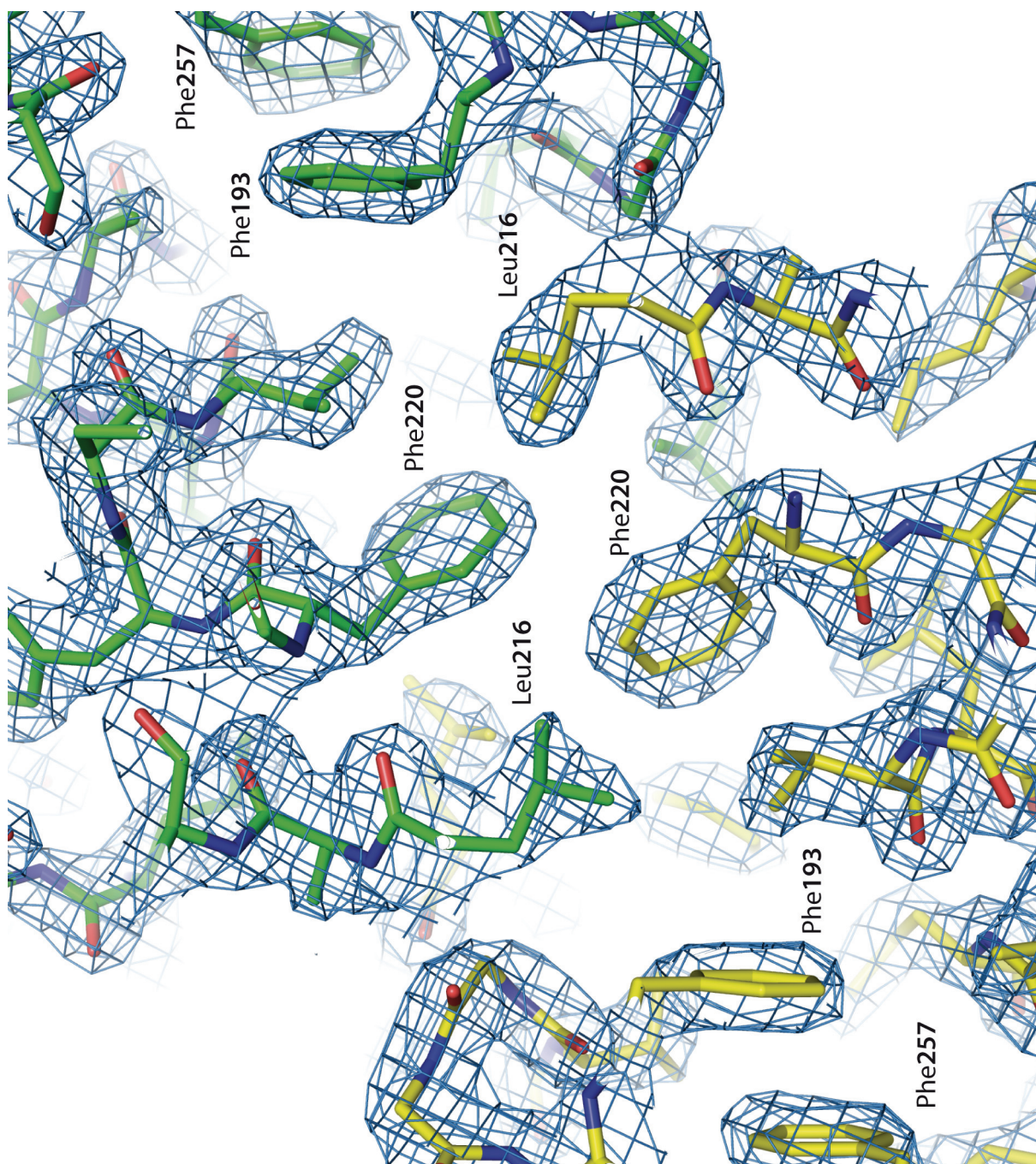




**Figure 7. Role of Clusters 1 and 2 in tRNA<sup>Tyr</sup> Acceptor Arm Recognition**

(A) Superimposition of the cleft formed by the two helical structures of clusters 1 and 2 (in which binds the tyrosine acceptor arm of tRNA<sup>Tyr</sup>) in the crystallographic structures of human mt-TyrRS (in brown), *B. stearothermophilus* TyrRS (in blue) and *T. thermophilus* TyrRS in complex with tRNA<sup>Tyr</sup> (in green). Notice the quasi-perfect superimposition of the two clusters in human mt-TyrRS and *B. stearothermophilus* TyrRS and the important structural deviations in *T. thermophilus* TyrRS. The bar at the bottom of the figure shows the position where the cleft is largest ( $d = 9.9, 10.0$  and  $12.4 \text{ \AA}$  in the TyrRSs from human mitochondria, *B. stearothermophilus* and *T. thermophilus*, respectively, indicating an enlargement of the cleft in *T. thermophilus* of  $\sim 2.5 \text{ \AA}$ ). The three amino acids that were mutagenized are indicated.

(B) View of the clusters and their proximity with the tRNA acceptor branch as seen in the crystal structure of the *T. thermophilus* complex (Yaremchuk et al., 2002).

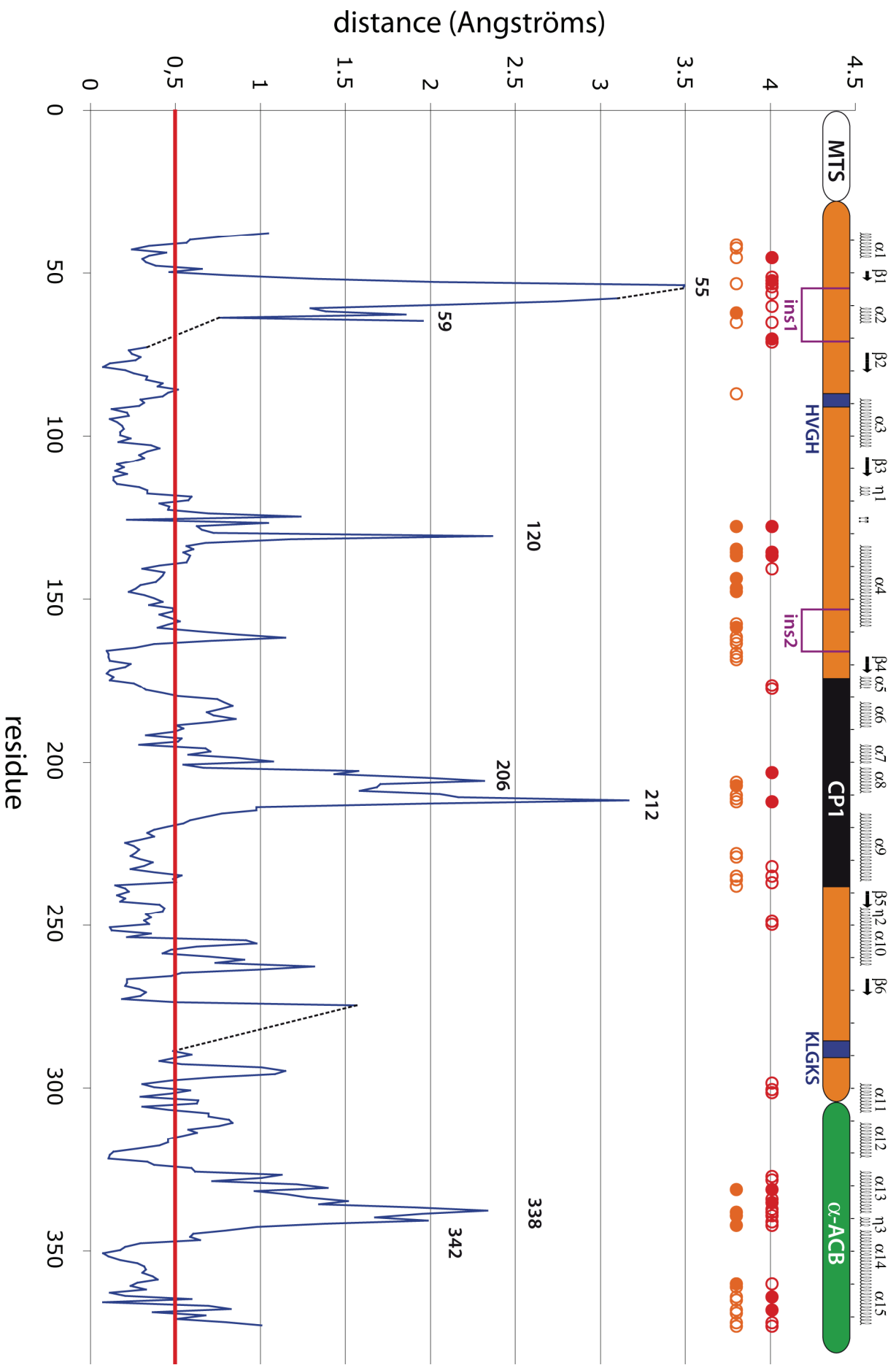


### Supplemental Data 1.

Example of the electron density map (contoured at  $1.5 \sigma$ ) showing the quality of the experimental data. The close-up view displays a detail of the interface between the two subunits made by residues of the CP1 domains.

### Supplemental Data 2.

$C_{\alpha}$ - $C_{\alpha}$  distances calculated after least squares minimized superimposition of two subunits of mt-TyrRS- $\Delta S4$  in complex with Tyr-AMS. The red line indicates the overall r.m.s.d. Crystallographic contacts for each monomer (orange for monomer A, red for monomer B) are indicated by dots above the graphic (full dots for contacts  $< 3.2 \text{ \AA}$ ; empty dots for contacts in the  $3.2\text{--}4.0 \text{ \AA}$  range). To facilitate data interpretation, the figure recalls the modular organization of the TyrRS (color code as in Figure 1A) with structural details (as in Figure 1B).











366 Cenorhabditis briggsae-mito  
 367 Cenorhabditis elegans-mito  
 368 Gallus gallus-mito  
 369 Drosophila melanogaster-mito  
 370 Canis familiaris-mito  
 371 Felis catus-mito  
 372 Bos taurus-mito  
 373 Sus scrofa-mito  
 374 Homo sapiens-mito  
 375 Mus musculus-mito  
 376 Rattus norvegicus-mito  
 377 Danio rerio-mito  
 378 Oryzias latipes-mito  
 379 Caenorhabditis elegans-mito  
 380 Gastrostacus aculeatus-mito  
 381 Takiyagu rubripes-mito  
 382 Ciona savignyi-mito  
 383 Ciona intestinalis-mito  
 384 Drosophila melanogaster-mito  
 385 Drosophila albomicans-mito  
 386 Drosophila yakuba-mito  
 387 Anopheles gambiae-mito  
 388 Arabidopsis thaliana-mito  
 389 Populus trichocarpa-mito  
 390 Neurospora crassa-mito  
 391 Schistosoma japonicum-mito  
 392 Schistosoma mansoni-mito  
 393 Candida albicans-mito  
 394 Debaryomyces hansenii-mito  
 395 Kluyveromyces fragilis-mito  
 396 Kluyveromyces fragilis-mito  
 397 Kluyveromyces fragilis-mito  
 398 Kluyveromyces fragilis-mito  
 399 Kluyveromyces fragilis-mito  
 400 Kluyveromyces fragilis-mito  
 401 Kluyveromyces fragilis-mito  
 402 Kluyveromyces fragilis-mito  
 403 Kluyveromyces fragilis-mito  
 404 Kluyveromyces fragilis-mito  
 405 Kluyveromyces fragilis-mito  
 406 Kluyveromyces fragilis-mito  
 407 Kluyveromyces fragilis-mito  
 408 Kluyveromyces fragilis-mito  
 409 Kluyveromyces fragilis-mito  
 410 Kluyveromyces fragilis-mito  
 411 Kluyveromyces fragilis-mito  
 412 Kluyveromyces fragilis-mito  
 413 Kluyveromyces fragilis-mito  
 414 Kluyveromyces fragilis-mito  
 415 Kluyveromyces fragilis-mito  
 416 Kluyveromyces fragilis-mito  
 417 Kluyveromyces fragilis-mito  
 418 Kluyveromyces fragilis-mito  
 419 Kluyveromyces fragilis-mito  
 420 Kluyveromyces fragilis-mito  
 421 Kluyveromyces fragilis-mito  
 422 Kluyveromyces fragilis-mito  
 423 Kluyveromyces fragilis-mito  
 424 Kluyveromyces fragilis-mito  
 425 Kluyveromyces fragilis-mito  
 426 Kluyveromyces fragilis-mito  
 427 Kluyveromyces fragilis-mito  
 428 Kluyveromyces fragilis-mito  
 429 Kluyveromyces fragilis-mito  
 430 Kluyveromyces fragilis-mito  
 431 Kluyveromyces fragilis-mito  
 432 Kluyveromyces fragilis-mito  
 433 Kluyveromyces fragilis-mito  
 434 Kluyveromyces fragilis-mito  
 435 Kluyveromyces fragilis-mito  
 436 Kluyveromyces fragilis-mito  
 437 Kluyveromyces fragilis-mito  
 438 Kluyveromyces fragilis-mito  
 439 Kluyveromyces fragilis-mito  
 440 Kluyveromyces fragilis-mito  
 441 Kluyveromyces fragilis-mito  
 442 Kluyveromyces fragilis-mito  
 443 Kluyveromyces fragilis-mito  
 444 Kluyveromyces fragilis-mito  
 445 Kluyveromyces fragilis-mito  
 446 Kluyveromyces fragilis-mito  
 447 Kluyveromyces fragilis-mito  
 448 Kluyveromyces fragilis-mito  
 449 Kluyveromyces fragilis-mito  
 450 Kluyveromyces fragilis-mito  
 451 Kluyveromyces fragilis-mito  
 452 Kluyveromyces fragilis-mito  
 453 Kluyveromyces fragilis-mito  
 454 Kluyveromyces fragilis-mito  
 455 Kluyveromyces fragilis-mito  
 456 Kluyveromyces fragilis-mito  
 457 Kluyveromyces fragilis-mito  
 458 Kluyveromyces fragilis-mito  
 459 Kluyveromyces fragilis-mito  
 460 Kluyveromyces fragilis-mito  
 461 Kluyveromyces fragilis-mito  
 462 Kluyveromyces fragilis-mito  
 463 Kluyveromyces fragilis-mito  
 464 Kluyveromyces fragilis-mito  
 465 Kluyveromyces fragilis-mito  
 466 Kluyveromyces fragilis-mito  
 467 Kluyveromyces fragilis-mito  
 468 Kluyveromyces fragilis-mito  
 469 Kluyveromyces fragilis-mito  
 470 Kluyveromyces fragilis-mito  
 471 Kluyveromyces fragilis-mito  
 472 Kluyveromyces fragilis-mito  
 473 Kluyveromyces fragilis-mito  
 474 Kluyveromyces fragilis-mito  
 475 Kluyveromyces fragilis-mito  
 476 Kluyveromyces fragilis-mito  
 477 Kluyveromyces fragilis-mito  
 478 Kluyveromyces fragilis-mito  
 479 Kluyveromyces fragilis-mito  
 480 Kluyveromyces fragilis-mito  
 481 Kluyveromyces fragilis-mito  
 482 Kluyveromyces fragilis-mito  
 483 Kluyveromyces fragilis-mito  
 484 Kluyveromyces fragilis-mito  
 485 Kluyveromyces fragilis-mito  
 486 Kluyveromyces fragilis-mito  
 487 Kluyveromyces fragilis-mito  
 488 Kluyveromyces fragilis-mito  
 489 Kluyveromyces fragilis-mito  
 490 Kluyveromyces fragilis-mito  
 491 Kluyveromyces fragilis-mito  
 492 Kluyveromyces fragilis-mito  
 493 Kluyveromyces fragilis-mito  
 494 Kluyveromyces fragilis-mito  
 495 Kluyveromyces fragilis-mito  
 496 Kluyveromyces fragilis-mito  
 497 Kluyveromyces fragilis-mito  
 498 Kluyveromyces fragilis-mito  
 499 Kluyveromyces fragilis-mito  
 500 Kluyveromyces fragilis-mito

446 Cenorhabditis briggsae-mito  
 447 Cenorhabditis elegans-mito  
 448 Gallus gallus-mito  
 449 Drosophila melanogaster-mito  
 450 Monodelphis domestica-mito  
 451 Canis familiaris-mito  
 452 Felis catus-mito  
 453 Bos taurus-mito  
 454 Sus scrofa-mito  
 455 Homo sapiens-mito  
 456 Mus musculus-mito  
 457 Rattus norvegicus-mito  
 458 Danio rerio-mito  
 459 Oryzias latipes-mito  
 460 Caenorhabditis elegans-mito  
 461 Takiyagu rubripes-mito  
 462 Ciona savignyi-mito  
 463 Ciona intestinalis-mito  
 464 Drosophila melanogaster-mito  
 465 Drosophila albomicans-mito  
 466 Drosophila yakuba-mito  
 467 Anopheles gambiae-mito  
 468 Arabidopsis thaliana-mito  
 469 Populus trichocarpa-mito  
 470 Neurospora crassa-mito  
 471 Schistosoma japonicum-mito  
 472 Schistosoma mansoni-mito  
 473 Candida albicans-mito  
 474 Debaryomyces hansenii-mito  
 475 Kluyveromyces fragilis-mito  
 476 Kluyveromyces fragilis-mito  
 477 Kluyveromyces fragilis-mito  
 478 Kluyveromyces fragilis-mito  
 479 Kluyveromyces fragilis-mito  
 480 Kluyveromyces fragilis-mito  
 481 Kluyveromyces fragilis-mito  
 482 Kluyveromyces fragilis-mito  
 483 Kluyveromyces fragilis-mito  
 484 Kluyveromyces fragilis-mito  
 485 Kluyveromyces fragilis-mito  
 486 Kluyveromyces fragilis-mito  
 487 Kluyveromyces fragilis-mito  
 488 Kluyveromyces fragilis-mito  
 489 Kluyveromyces fragilis-mito  
 490 Kluyveromyces fragilis-mito  
 491 Kluyveromyces fragilis-mito  
 492 Kluyveromyces fragilis-mito  
 493 Kluyveromyces fragilis-mito  
 494 Kluyveromyces fragilis-mito  
 495 Kluyveromyces fragilis-mito  
 496 Kluyveromyces fragilis-mito  
 497 Kluyveromyces fragilis-mito  
 498 Kluyveromyces fragilis-mito  
 499 Kluyveromyces fragilis-mito  
 500 Kluyveromyces fragilis-mito

Les informations apportées par cette structure sont nombreuses, mais ne répondent pas à toutes les questions soulevées par les tests fonctionnels. Le mécanisme précis de reconnaissance de l'extrémité acceptrice de l'ARNt<sup>Tyr</sup> mitochondrial par la mt-TyrRS reste notamment à élucider. Ceci pourrait être fait en étendant les expériences de mutagenèse réalisées au niveau des clusters 1 et 2 aux acides aminés potentiellement en contact avec l'ARNt<sup>Tyr</sup>. Citons l'Arg201 du cluster1, la Leu248 du cluster2 ou l'Asp311 du domaine de liaison à l'anticodon. Il serait aussi très important de déterminer la structure du complexe mt-TyrRS/mt-ARNt<sup>Tyr</sup>. Une telle structure révélerait aussi le mécanisme de fixation de la tyrosine sur l'extrémité 3' terminale de l'ARNt<sup>Tyr</sup>. Chez la TrpRS, la seule autre synthétase de classe I avec la TyrRS à présenter une organisation dimérique et un mode de fixation de l'ARNt caractéristique des aaRS de classe II, le mode de reconnaissance et le changement de conformation de la séquence CCA 3' terminal de l'ARNt<sup>Trp</sup> induit pour l'aminoacylation sont caractéristiques des aaRS de classe I.

Connaissant la proximité évolutive et structurale entre TrpRS et TyrRS (Ribas de Pouplana et al., 1996), il est probable que la TyrRS utilise le même mécanisme. Néanmoins, dans les structures de complexes ARNt<sup>Tyr</sup>/TyrRS résolues à ce jour cette extrémité CCA est soit désordonnée, soit en dehors du site actif. De la même façon, le mode d'interaction au niveau de l'anticodon n'est pas encore compris. Les alignements de séquences n'ont pour le moment permis d'identifier qu'un seul résidu susceptible de contacter la base 34 de l'anticodon, alors que dans le cas de la mitochondrie humaine, les 3 positions du triplet anticodon ont été montrées ayant un rôle fort dans l'identité. Les tentatives de positionnement de l'ARNt<sup>Tyr</sup> eubactérien de *T. thermophilus* (Yaremchuk et al., 2002) sur la TyrRS mitochondriale se révèlent toutes aussi inefficaces à proposer un modèle cohérent d'interaction entre le triplet anticodon et les résidus de la mt-TyrRS. Il est probable que la conformation et la flexibilité de l'ARNt<sup>Tyr</sup> mitochondrial, notamment de son triplet anticodon permettent de réaliser des contacts bases spécifiques avec la synthétase.

Il serait aussi intéressant de déterminer la structure du domaine C-terminal S4-like seul ou dans le cadre du complexe mt-tRNA<sup>Tyr</sup>/mt-TyrRS. En effet, ce domaine interagit chez les TyrRS eubactériennes avec l'anticodon et la grande région variable de l'ARNt<sup>Tyr</sup> homologue. Ce domaine est conservé chez les TyrRS mitochondriales, mais la grande région variable a disparu chez les ARNt<sup>Tyr</sup> mitochondriaux de vertébrés. La question du rôle de ce domaine dans ces systèmes est donc sujette à questions. Un premier élément de réponse est apporté par les propriétés d'aminoacylation de la mt-TyrRS humaine privée de ce domaine. Cette dernière a perdu la capacité d'aminoacyler les transcrits d'ARNt<sup>Tyr</sup> de *E. coli*, de *S. cerevisiae* et humains, mais aussi l'ARNt<sup>Tyr</sup> natif de *S. cerevisiae*. Seul l'ARNt<sup>Tyr</sup> natif

d'*E. coli* est encore chargé par cette enzyme, mais avec une perte d'efficacité de 100 fois par rapport à l'enzyme entière, principalement due à un effet  $k_{cat}$ . Cet effet suggère un rôle de stabilisation de l'ARNt<sup>Tyr</sup> plutôt qu'une reconnaissance spécifique pour le domaine S4-like dans le cas de la mitochondrie humaine. Il permettrait ainsi de compenser l'absence de modifications post-transcriptionnelles des transcrits testés. Pourtant l'ajout de ce domaine libre dans une réaction d'aminacylation d'ARNt<sup>Tyr</sup> d'*E. coli* par la TyrRS mitochondriale humaine réduit l'efficacité d'aminacylation. L'efficacité est encore plus réduite si la mt-TyrRS testée ne possède plus le domaine S4-like. Ceci suggère une compétition entre la synthétase et le domaine libre pour la fixation de l'ARNt<sup>Tyr</sup>. Ces résultats contradictoires ne permettent donc pas de définir clairement le rôle de ce domaine dans la réaction de tyrosylation au niveau de la mitochondrie humaine.

### 3. Cystéines, ponts disulfure et activité

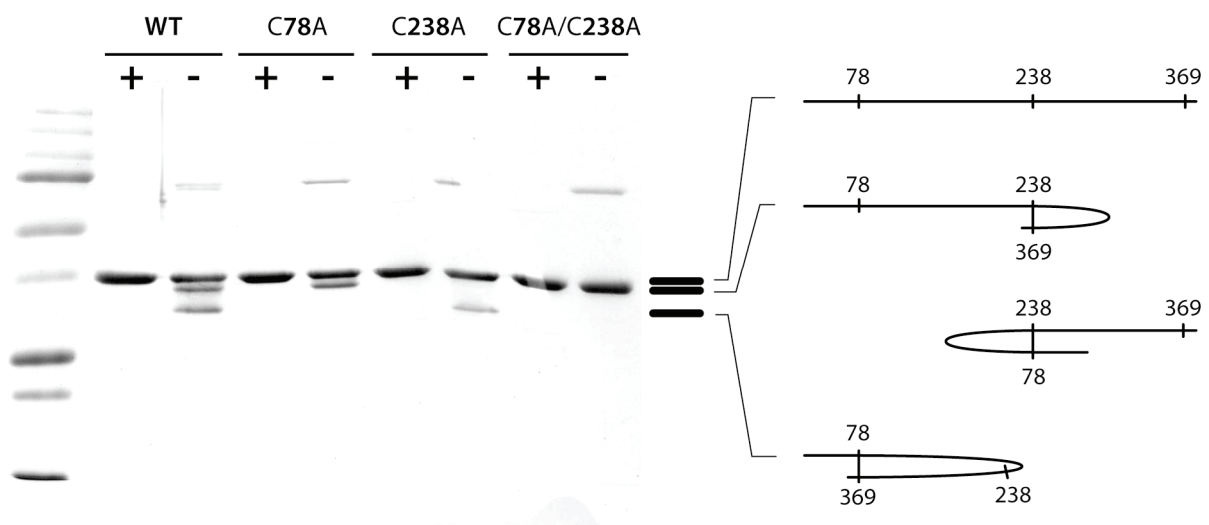
Au cours du travail de caractérisation fonctionnelle et structurale de la mt-TyrRS et de la mt-TyrRS- $\Delta$ S4 nous nous sommes interrogés sur l'état d'oxydoréduction de cette enzyme. D'une part, les réactions métaboliques de la chaîne respiratoire font de la matrice mitochondriale un environnement particulier, longtemps considéré comme oxydant, mais récemment démontré comme étant réducteur (Korge et Weiss, 2006). D'autre part, les cystéines présentes dans la TyrRS de *S. cerevisiae* ont été proposées comme participant au mécanisme catalytique de l'enzyme (Faulhammer et Cramer, 1977). En ce qui concerne la mt-TyrRS, chaque monomère comporte 4 cystéines dans sa séquence (Cys78 et Cys238 dans le domaine catalytique, Cys369 à l'extrémité du domaine de liaison à l'anticodon, et la dernière Cys413 dans le domaine C-terminal S4-like). Ce nombre est donc réduit à 3 pour la mt-TyrRS- $\Delta$ S4. Au regard de leur localisation dans la structure, leur implication dans le mécanisme réactionnel est peu probable (**figure 2** de l'**article 5**). De même, la formation d'un pont disulfure semble impossible sans perturber fortement la structure intrinsèque de l'enzyme, les deux cystéines les plus proches dans le domaine catalytique étant éloignées de 10 Å. Néanmoins, nous avons observé très tôt dans l'étude de la mt-TyrRS la présence de deux bandes sur gel de protéines (Résultat 4 de la **figure 13**) que le séquençage N-terminal avait identifiées comme étant la mt-TyrRS intègre. L'intégrité à l'extrémité C-terminale est confirmée grâce à la queue poly-histidine fixée à celle-ci. En effet, toute dégradation du côté C-terminal aurait empêché la protéine de se fixer sur la résine substituée avec du Nickel. Toute possibilité de dégradation de l'enzyme, qui aurait pu expliquer la présence de bandes

supplémentaires, est donc exclue. Ce même phénomène s'est révélé avec la mt-TyrRS- $\Delta$ S4, pour laquelle deux bandes supplémentaires apparaissent sur gel en conditions non réductrices (**figure 23**). Comme pour la mt-TyrRS complète, les bandes surnuméraires et de tailles apparentes inférieures à celle du monomère correspondant disparaissent sur gel en conditions réductrices. Tout indique donc que ces bandes surnuméraires correspondent à des formes oxydées de l'enzyme pour la mt-TyrRS et la mt-TyrRS- $\Delta$ S4.

Pour pouvoir isoler l'une ou l'autre des formes nous avons traité les protéines avec le réactif d'Ellman ou DTNB (5,5'-dithio-bis(2-nitrobenzoic acid) ; Ellman, 1959). Ce réactif, classiquement utilisé pour le dosage des cystéines, peut aussi être employé pour induire la formation de ponts disulfure (pour stabiliser la structure de protéines; Ivens et al., 2002). Dans notre cas nous avons rajouté une étape de lavage supplémentaire au protocole de purification par chromatographie d'affinité : après fixation de la mt-TyrRS sur la résine, celle-ci est lavée avec 120 mL de tampon contenant 5 mM de DTNB. Le reste de la purification se poursuit ensuite de manière classique sans agent oxydant. Les différentes formes isolées (après vérification sur gel de protéines) ont été testées fonctionnellement dans le cas de la mt-TyrRS complète et cristallisées dans le cas de la mt-TyrRS- $\Delta$ S4. La mt-TyrRS oxydée est toujours capable d'activer la tyrosine mais a perdu la capacité de le transférer sur l'ARNt, ce qui indique que la perturbation engendrée n'affecte pas la conformation du site catalytique et que la forme active de l'enzyme correspond à la forme réduite. La mt-TyrRS- $\Delta$ S4 a pu être cristallisée sous forme oxydée et réduite. Les cristaux de la forme oxydée, probablement moins stables ne sont pas de qualité suffisante pour permettre d'enregistrer des clichés de diffraction exploitables. Les cristaux de la forme réduite par contre diffractent les rayons X à plus de 3 Å de résolution.

Afin de comprendre le phénomène d'oxydoréduction auquel la mt-TyrRS est sujette en solution nous avons entrepris de muter les résidus cystéines de la forme tronquée, mt-TyrRS- $\Delta$ S4. Les deux résidus cystéines du domaine catalytique ont été mutés en alanine, indépendamment et conjointement. Les différents mutants sont tous actifs en aminoacylation, malgré une légère perte d'activité, ce qui confirme la non implication des cystéines dans le mécanisme réactionnel. Les profils de migration sur gel de polyacrylamide dénaturant des différents mutants permettent d'identifier la nature du pont disulfure formé pour chaque bande sur le gel (**figure 23**). Notons aussi l'absence d'une bande pour l'enzyme non réduite qui correspondrait à la formation d'un pont disulfure entre les deux cystéines du domaine catalytique. Ceci est probablement dû à la stabilité du domaine catalytique qui est trop structuré pour permettre le rapprochement de ces deux résidus (éloignés de 10 Å, ils devraient se rapprocher de 8 Å pour former un pont disulfure). La

cystéine 369 par contre est à l'extrémité du monomère et peut plus facilement venir interagir avec l'une des deux autres cystéines restantes. Enfin, le double mutant C78A/C238A de la mt-TyrRS- $\Delta$ S4 a été soumis au traitement oxydant et son activité testée. Cette enzyme oxydée est toujours capable d'aminocycler l'ARNt<sup>Tyr</sup> d'*E. coli* avec la tyrosine, ce qui prouve que le traitement infligé à la protéine cible spécifiquement les cystéines et que la perte d'activité observée sur la mt-TyrRS complète oxydée résulte bien du blocage de l'enzyme dans une conformation inactive. Enfin, la présence de seulement deux bandes pour cette mt-TyrRS en conditions non réductrices (**figure 13**) malgré la présence de 4 cystéines par monomère peut s'expliquer par la formation préférentielle d'un seul pont disulfure entre la cystéine 369 du domaine de liaison à l'anticodon et la cystéine 413 du domaine C-terminal S4-like, plutôt qu'avec une des cystéines du domaine catalytique. Cette hypothèse est en accord avec la taille apparente de la deuxième bande sur le gel de polyacrylamide. En conclusion, la formation de ponts disulfures dans la structure de la TyrRS mitochondriale n'est pas un phénomène qui participe à son activité catalytique mais qui semble plutôt refléter la plasticité de cette TyrRS, particulièrement entre ses différents domaines.



**Figure 23. Implication des ponts disulfures dans la conformation de la mt-TyrRS- $\Delta$ S4**

La conformation de la mt-TyrRS- $\Delta$ S4 sauvage (WT) ou mutée sur un ou deux de ses résidus cystéines (C78A, cystéine78 en alanine ; C238A ; double mutant C78A/C238A) est analysée sur gel de polyacrylamide dénaturant 10% avec un tampon de dépôt avec (+) ou sans (-) agent réducteur. La conformation correspondant à chaque bande visualisée sur gel est indiquée sur la droite de la figure. La numérotation des résidus est celle utilisée dans les articles et retrouvée dans les banques de données. Le résidu N-terminal de la mt-TyrRS- $\Delta$ S4 est l'acide aminé 28.

## D. Les TyrRS mitochondriales à travers l'évolution

Afin de dégager des caractéristiques communes à toutes ou certaines classes de TyrRS mitochondriales, la comparaison de leurs séquences chez des organismes eucaryotes variés est nécessaire. Malheureusement, à l'heure actuelle, seules 6 séquences de TyrRS mitochondriales sont clairement identifiées, il s'agit de celles des levures *N. crassa*, *Podospora anserina* et *S. cerevisiae*, de la souris, de la drosophile et de l'Homme. Ce nombre est insuffisant et l'échantillon n'est pas assez représentatif des différents groupes d'eucaryotes.

### 1. Recherche de séquences de TyrRS mitochondriales

Pour pouvoir réaliser une classification fiable des TyrRS mitochondriales, nous avons recherché de nouvelles séquences dans les banques de données accessibles soit *via* Ensembl (<http://www.ensembl.org/>) soit *via* les sites dédiés au séquençage de chaque organisme. Les candidats identifiés ont été validés par la présence d'un signal d'adressage vers la mitochondrie (prédiction avec l'algorithme MITOPROT ; Claros et Vincens, 1996) et des alignements avec les autres TyrRS mitochondriales connues et identifiées. Les séquences peptidiques partielles ont été, le cas échéant, ajustées et assemblées sur la base de l'alignement avec les séquences de TyrRS mitochondriales déjà disponibles. Cette méthode a ainsi permis d'identifier la séquence de TyrRS mitochondriales de plusieurs organismes en cours de séquençage. Les séquences reconstruites se sont révélées identiques à celles disponibles après séquençage complet des organismes. Cette recherche a permis d'obtenir une cinquantaine de séquences de TyrRS mitochondriales de provenances variées : 13 levures, 3 plantes, 5 insectes, 5 poissons, 2 nématodes, 2 cnidaires et 21 animaux dont 15 mammifères. Leurs séquences sont consultables en **Annexes**. Les mammifères, qui constituent la forme la plus évoluée d'eucaryotes, sont très étudiés ce qui explique le nombre important de séquences identifiées chez ces organismes, comparativement à l'extrême diversité d'eucaryotes existants.



<b>Organisme</b>	<b>Nom commun</b>	<b>Taille (aa)</b>	<b>Probabilité d'adressage</b>	<b>Premier aa après MTS</b>	<b>Acces code</b>	<b>locus</b>
<i>Aedes aegypti</i>	Moustique	463	0,9604	14	AAEL010563-PA	
<i>Anopheles gambiae</i>	Anophèle	462	0,9893	21	ENSANGP00000004765	
<i>Apis mellifera</i>	Abeille	457	0,9773	18	XP_624409	
<i>Arabidopsis thaliana</i>		511	0,9769	47	NP_186915	
<i>Bos taurus</i>	Vache	478	0,9276	17	ENSBTAP00000015149	
<i>Caenorhabditis briggsae</i>		455	0,6695	21	ENSCBRP00000005606	
<i>Caenorhabditis elegans</i>		447	0,8769	21	AAB09162.2	K08F11.4a
<i>Candida albicans</i>		501	0,9458	26	EAK95246	
<i>Candida glabrata</i>		462	0,9882	23		CAG59980
<i>Canis familiaris</i>	Chien	477	0,9785	34	XP_543740	
<i>Cavia Porcellus</i>	Cochon d'Inde	477	0,9766	?	ENSCPOG00000013767	
<i>Ciona intestinalis</i>	Corail	514	0,9716	46	ENSCINP00000015337	
<i>Ciona savignyi</i>	Eponge	553	0,993	84	GENSCAN00000062141	
<i>Danio rerio</i>	Poisson Zèbre	485	0,9951	21	BC055576/Q7SXJ3	
<i>Dasypus novemcinctus</i>	Tatou	200	-	-		
<i>Debaryomyces hansenii</i>		523	0,6716	29		CAG88504
<i>Drosophila melanogaster</i>	Drosophile	464	0,9298	23	Q9W107	SYYM_DROME
<i>Drosophila pseudoobscura</i>		464	0,9889	31	EAL25794	EAL25794
<i>Echinops telfairi</i>	Hérisson	200	-	-		
<i>Erinaceus europaeus</i>	Hérisson	364	-	-	ENSEEUP00000007475	
<i>Felis catus</i>	Chat	477	0,8672	23	ENSFCAP00000004942	
<i>Gallus gallus</i>	Poule	484	0,9971	39		
<i>Gasterosteus aculeatus</i>	Epinoche	472	0,9951	16	ENSGACP00000004911	
<i>Homo sapiens</i>		477	0,974	17	Q9Y2Z4	SYYM_HUMAN
<i>Kluyberomyces lactis</i>		453	0,9064	19		CAH00131
<i>Loxodonta africana</i>	Elephant	276	-	-	ENSLAFP00000014557	
<i>Macaca mulatta</i>	Macaque	316	0,9546	17		
<i>Monodelphis domestica</i>	Opossum	471	0,9944	29	From_Q9Y2Z4_etc_1	
<i>Mus Musculus</i>	Souris	472	0,9995	46	ENSMUSP00000055277	SYYM_MOUSE
<i>Neurospora crassa</i>		669	0,9895	45	P12063	SYYM_NEUCR
<i>Oriza sativa</i>	Riz	497	0,9977	49	NP_918582	
<i>Ornithorhynchus anatinus</i>	Ornithorhynque	220	-	-		
<i>Oryctolagus cuniculus</i>	Lapin	478	0,9936	15	ENSOCUG00000000022	
<i>Oryzias latipes</i>	Médaka	473	0,9943	45	ENSORLP00000003010	
<i>Pan troglodytes</i>	Chimpanzé	477	0,974	17	ENSPTRP00000008235	
<i>Podospira anserina</i>		640	0,8646	?	P28669	SYYM_PODAN
<i>Populus trichocarpa</i>	Peuplier	439	-	-		
<i>Ratus norvegicus</i>	Rat	471	0,998	46	ENSRNOP00000037398	NP_001009627
<i>Saccharomyces bayanus</i>		493	0,9609	38		
<i>Saccharomyces castellii</i>		459	0,9329	37		
<i>Saccharomyces cerevisiae</i>		492	0,8476	38	P48527	SYYM_YEAST
<i>Saccharomyces kudriavzevii</i>		492	0,9161	37		
<i>Saccharomyces mitakae</i>		493	0,9012	38		
<i>Schizosaccharomyces pombe</i>		445	0,7418	13	NP_588433	
<i>Strongylocentrotus purpuratus</i>	Oursin	560	0,1067	29	XP_781450	
<i>Sus scrofa</i>	Cochon	478	0,9886	23		
<i>Takifugu rubripes</i>	Fugu	459	-	-	SINFRUP00000153575	
<i>Tetodon nigroviridis</i>	Poisson coffre	480	0,9885	18	GSTENT00011859001	
<i>Tupaia belangeri</i>	Tupaia	478	0,9919	45	ENSTBEP00000007161	
<i>Xenopus tropicalis</i>	Xénope	354	0,9964	31	ENSXETP00000026810	
<i>Yarovia lypolytica</i>		527	0,9843	29	XP_504538	CAG80141

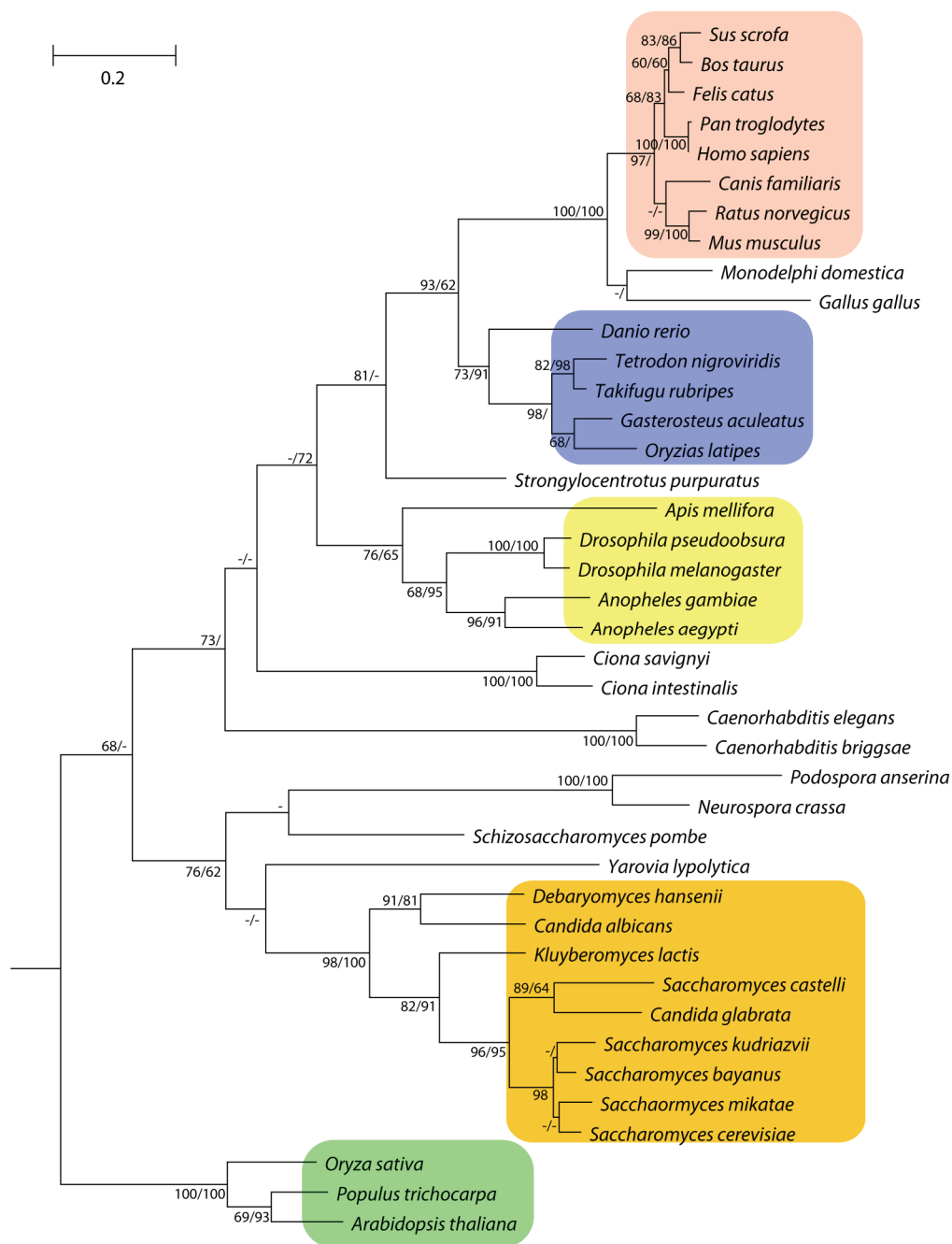
**Tableau 7. Séquences de TyrRS mitochondriales**

Les séquences intègres de TyrRS mitochondriales ont été utilisées pour réaliser un alignement multiple (matériel supplémentaire n°5 de l'**article n°5**). On y retrouve une très forte conservation du domaine catalytique où se localise la majorité des résidus strictement conservés, dont ceux en contact direct avec la tyrosine et l'ATP. Ainsi le mode de fixation de ces deux substrats par la TyrRS semble bien être conservé chez toutes les espèces. Les zones les plus variables de ce domaine se situent à son extrémité N-terminal et plus précisément au niveau de l'insertion "ins1", mais aussi au niveau des insertions "ins2" et "II". Le domaine de liaison à l'anticodon, dont la taille varie très peu, ne comprend que 4 résidus strictement conservés. Parmi eux, le résidu Asp311 a été proposé comme impliqué dans la reconnaissance de la base G34 de l'anticodon, mais les 3 autres acides aminés sont trop éloignés du site de fixation probable de l'anticodon pour participer à sa reconnaissance. Enfin, le domaine C-terminal, homologue à la protéine ribosomale S4 présente le plus de variabilité dans sa taille et sa séquence. D'après la structure du complexe ARN<sup>Tyr</sup>/TyrRS de *T. thermophilus* (Yaremchuk et al., 2002) les zones les plus conservées de ce domaine chez les TyrRS mitochondriales correspondraient aux zones de contact avec la branche anticodon de l'ARN<sup>Tyr</sup>.

## 2. Analyse phylogénétique

Le nombre important de séquences de TyrRS mitochondriales rassemblées ainsi que leur large répartition parmi les différents groupes d'eucaryotes permettant de construire un arbre phylogénétique. Pour ce faire l'alignement a été restreint au seul domaine catalytique. De même, toutes les insertions ont été supprimées afin d'éviter tout biais. L'arbre a été construit avec deux méthodes différentes, celle des plus proches voisins (Neighbor Joining, NJ) et celle du maximum de vraisemblance (Maximum Likelihood, ML).

On retrouve à travers cet arbre la classification générale des différents groupes d'eucaryotes. Ces différents groupes (levures, plantes, poissons, mammifères) sont clairement séparés les uns des autres. On retrouve comme pour les TyrRS cytoplasmiques une très grande distance évolutive entre mt-TyrRS d'animaux et de plantes. Ceci reflète le codage nucléaire des TyrRS mitochondriales. Seules les mt-TyrRS des levures *N. crassa* et *P. anserina* et celles de nématodes se démarquent par leur positions isolées. La position des deux dernières peut s'expliquer par leur fonction additionnelle d'épissage. La position des mt-TyrRS de nématodes est par contre plus surprenante et appelle une étude plus poussée de leur séquence.



**Figure 24. Arbre phylogénétique des TyrRS mitochondriales.**

Réalisé suivant le même principe que l'arbre des TyrRS présenté en figure 10. Un fond de couleur souligne les différents groupes : rose pour les mammifères, bleu pour les poissons, jaune pour les insectes, orange pour les levures et vert pour les plantes. L'arbre est enraciné avec la séquence de la TyrRS d'*E. coli* comme "out-group". Les valeurs de bootstap supérieures à 60% pour les méthodes NJ et ML sont indiquées, respectivement.

# Conclusions et Perspectives



## **Jeu d'identité du mt-ARNt<sup>Tyr</sup> et mt-TyrRS non spécifique**

Le travail réalisé au cours de ma thèse a visé à caractériser une tyrosyl-ARNt synthétase phylogénétiquement hybride, la TyrRS mitochondriale humaine. Ce choix a été fait dans le but de comprendre l'évolution des systèmes de tyrosylation dans leur ensemble, mais aussi plus particulièrement au niveau de la "barrière d'espèces" qui sépare systèmes eubactériens et archaea/eucaryotes. Nous avons, à partir de sa séquence primaire identifiée dans les banques de données, caractérisé cette enzyme par des approches complémentaires de biochimie, biologie moléculaire et structurale. L'identification du gène codant pour la TyrRS mitochondriale humaine a été faite dans une approche globale visant à caractériser l'ensemble du jeu de synthétases mitochondriales humaines (**articles n°1 et n°2**). L'analyse de la séquence protéique de la mt-TyrRS révélait déjà des particularités qui suggéraient un comportement atypique pour cette TyrRS par rapport aux TyrRS connues. Les tests fonctionnels ont confirmé cette hypothèse et ont montré que cette TyrRS est la première dépourvue de spécificité vis-à-vis de la paire de base N1-N72 des ARNt<sup>Tyr</sup> qui constitue la pierre angulaire de la "barrière d'espèces" (**article n°3**). Cette étude fonctionnelle a été complétée par la mise en évidence de 4 nucléotides de l'ARNt<sup>Tyr</sup> mitochondrial cruciaux pour la reconnaissance par la mt-TyrRS. Il s'agit de la base discriminatrice A73, ainsi que les 3 résidus du triplet anticodon (**article de revue n°1**). Chacun de ces nucléotides joue un rôle très fort comme pour compenser la perte de spécificité vis-à-vis de la première paire de base. Nous avons aussi analysé fonctionnellement les ARNt correspondant à trois mutations pathogéniques de l'ADNt mitochondrial humain situées dans les boucles D et T. Parmi ces 3 mutations, celle touchant le nucléotide en position 22 de la boucle D a un effet fort sur l'aminoacylation sans perturbation apparente de la structure de l'ARNt<sup>Tyr</sup> mitochondrial, ce qui en ferait un nouvel élément d'identité pour cet ARNt<sup>Tyr</sup>.

## **L'ARNt<sup>Tyr</sup> mitochondrial humain, jeu d'identité et structure**

La caractérisation du système ARNt<sup>Tyr</sup>/TyrRS dans la mitochondrie est loin d'être complète, mais plusieurs pistes sont ouvertes. Concernant l'ARNt<sup>Tyr</sup> tout d'abord, sa structure devra être déterminée, soit par cartographie en solution avec des sondes chimiques, soit par cristallographie. Cette structure devrait apporter des informations sur les modalités du repliement tertiaire en "L" des ARNt mitochondriaux. Des expériences d'empreintes de la mt-TyrRS sur l'ARNt<sup>Tyr</sup> mettront en évidence les résidus en contact avec la synthétase et

permettront de déterminer le mode d'interaction entre l'ARNt et la synthétase. Ces expériences pourraient aussi mettre en évidence des changements conformationnels de l'ARNt lors de l'interaction. Sur un autre plan, les expériences de transplantation du jeu d'identité devront être poursuivies pour pouvoir confirmer sa complétude.

### **Structure de la mt-TyrRS, classique et singulière à la fois**

Pour approfondir la caractérisation de la TyrRS mitochondriale humaine, son étude structurale par cristallographie a été initiée. Les problèmes rencontrés avec l'enzyme complète ont été résolus par la délétion du domaine C-terminal S4-like qui perturbait la cristallisation. Les cristaux obtenus ont permis d'atteindre une résolution de 2.7 Å (**article n°4**) puis de 2.2 Å en présence d'un analogue d'adénylate de tyrosine. La structure résolue montre une TyrRS mitochondriale homodimérique mais présentant une asymétrie entre ses deux monomères, reflet de son asymétrie fonctionnelle. Cette TyrRS est structurellement très proche des TyrRS de bactéries, telles que *E. coli* et *B. stearothermophilus*. Ainsi le mode de fixation du ligand est strictement conservé avec les TyrRS eubactériennes en particulier, mais également entre toutes les TyrRS. Par contre la TyrRS mitochondriale humaine présente deux insertions dans sa structure, absentes chez les homologues eubactériennes. La première insertion, d'une dizaine d'acides aminés, se situe à proximité du site de reconnaissance de l'extrémité acceptrice de l'ARNt<sup>Tyr</sup> mitochondrial et pourrait être impliquée dans la stabilisation de la branche acceptrice. De plus, cette insertion semble caractéristique des TyrRS mitochondriales de vertébrés. La seconde insertion, située sur la face "arrière" de la synthétase (la face interagissant avec l'ARNt étant considérée comme la face "avant") est plus importante, avec une vingtaine d'acides aminés, et retrouvée en séquence chez toutes les TyrRS mitochondriales. Le potentiel électrostatique de la synthétase révèle une forte proportion de charges positives sur chacune de ses deux faces, indice potentiel de son interaction avec des partenaires : ARNt, protéines ou membrane mitochondriale (**article n°5**). Enfin, trois résidus supposés importants pour la fixation de la branche acceptrice de l'ARNt<sup>Tyr</sup> mitochondrial ont été mutés. L'un d'eux a un effet dramatique sur la fixation de l'ARNt et interagit probablement avec la base discriminatrice A73 de l'ARNt<sup>Tyr</sup>. La mutation des deux autres acides aminés au contraire ne perturbe pas l'activité de la TyrRS (activation de l'acide aminé et transfert sur l'ARNt) ne permettant pas de définir clairement le mécanisme de fixation de l'extrémité acceptrice du mt-ARNt<sup>Tyr</sup>.



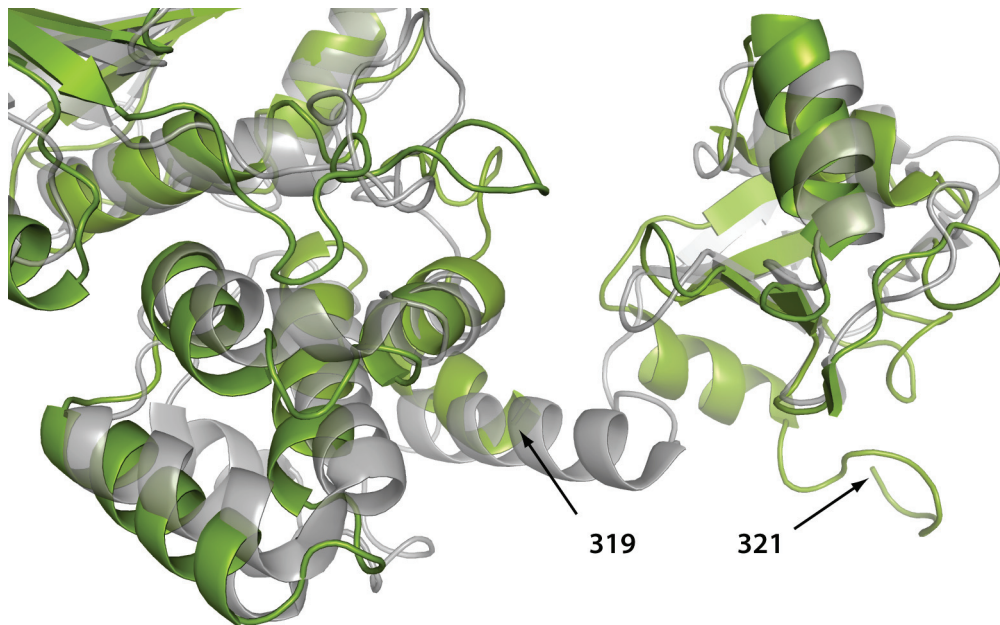
## La mt-TyrRS humaine, version longue, conformation et complexes

Du côté de la synthétase, un certain nombre de points restent à élucider. Pour comprendre les différentes étapes de la réaction d'aminoacylation, il serait utile de disposer de structures cristallographiques de l'enzyme complexée à un ou plusieurs substrats ou produits de la réaction d'aminoacylation qui formeraient une suite d'instantanés. Des essais de cristallisation ont déjà été faits avec ces différents substrats et devront être poursuivis. Plus particulièrement, la structure du complexe mt-ARNt<sup>Tyr</sup>/mt-TyrRS permettrait de comprendre le mécanisme d'interaction mis en place pour ce système et qui permet à l'enzyme d'adapter des ARNt<sup>Tyr</sup> hétérologues. Pour ce faire nous avons déjà tenté de co-cristalliser la mt-TyrRS complète ou tronquée avec l'ARNt<sup>Tyr</sup> natif d'*E. coli*. La formation d'un complexe avait été vérifiée par diffusion de lumière où une troisième population de molécules apparaissait en plus de l'aaRS et de l'ARNt seuls. Pour augmenter les chances de cristallisation de ce complexe une étape utile serait de le purifier par gel-filtration pour le débarrasser des composants isolés qui perturbent probablement sa cristallisation. La structure du complexe serait aussi un moyen de visualiser le domaine C-terminal S4-like qui devrait être stabilisé par la fixation de l'ARNt<sup>Tyr</sup>. Dans le cas contraire, la résolution de la structure isolée de ce domaine et sa comparaison aux domaines S4-like des TyrRS de *B. stearothermophilus* et de *T. thermophilus* devrait apporter des informations sur les contacts possibles avec l'ARNt<sup>Tyr</sup>. Parallèlement, la systématisation du processus de mutagenèse à tous les acides aminés potentiellement en interaction avec l'ARNt<sup>Tyr</sup> permettrait de confirmer le mode d'interaction avec la synthétase.

## Les TyrRS, vers une vision globale

Cette étude soulève aussi des questions plus générales sur les systèmes de tyrosylation, mais aussi sur les systèmes d'aminoacylation mitochondriaux. Les TyrRS sont probablement les synthétases les plus étudiées avec des structures résolues pour 4 bactéries, 4 archaea, 2 eucaryotes, 2 mitochondries et 1 virus. Parmi celles-ci, la structure de complexes ARNt<sup>Tyr</sup>/TyrRS est aussi connue chez une eubactérie, une archaea et un eucaryote. Toutes ces informations ont permis de comprendre la différence dans le mode de reconnaissance entre les deux groupes de TyrRS. Néanmoins certains points restent encore à élucider. On sait notamment que les TyrRS eubactériennes peuvent être séparées en deux groupes distincts, sans relation avec quelque critère taxonomique, mais suite à une duplication de gène survenue après la séparation des trois grands règnes du vivant. Cet état

de fait explique la similarité de 40% de la TyrRS mitochondriale humaine avec celles d'*E. coli* et *B. stearothermophilus* contre seulement 22% avec celle de *T. thermophilus*, qui fait elle partie du second groupe de TyrRS eubactériennes. De plus le domaine C-terminal S4-like est légèrement plus grand chez les TyrRS eubactériennes du premier groupe avec notamment une insertion au niveau de la liaison avec le domaine de liaison à l'anticodon.



**Figure 25. Superposition des structures de TyrRS eubactériennes avec domaine S4-like**

La structure de la TyrRS de *T. thermophilus* (pdb id 1H3E ; Yaremchuk et al., 2002) est représentée en gris transparent. La structure de la TyrRS minimale de *B. stearothermophilus* (pdb id 2TS1 ; Brick et al., 1989) et de son domaine S4-like (pdb id 1JH3 ; Guijarro et al., 2002) est représentée en vert. Le positionnement du domaine S4-like de *B. stearothermophilus* a été fait par superposition avec celui de *T. thermophilus*. Les numéros des résidus aux extrémités jointives du domaine de liaison à l'anticodon et S4-like sont indiqués.

Cette différence se voit d'autant mieux avec la structure de ce domaine pour les TyrRS eubactériennes cristallisées. Le domaine S4-like de *B. stearothermophilus* présente une première hélice absente chez *T. thermophilus*. De plus, quand on positionne ce domaine par rapport au domaine de liaison à l'anticodon chez *B. stearothermophilus* (le domaine S4-like ayant été cristallisé indépendamment) la liaison entre ces deux domaines n'est pas évidente et présentera des différences significatives par rapport à *T. thermophilus*. Il est donc probable que la structure d'un complexe ARNt<sup>Tyr</sup>/TyrRS de bactéries du premier groupe révèle des différences dans le mécanisme de reconnaissance de l'ARNt<sup>Tyr</sup> qui apporterait de nouveaux indices pour la compréhension de la mise en place de la "barrière d'espèces". Ces connaissances permettraient aussi de mettre au point des antibiotiques ciblant

spécifiquement les TyrRS eubactériennes sans interférer avec l'activité de la TyrRS mitochondriale.

### **Les synthétases mitochondriales, nouvelles règles du jeu ?**

Dans le domaine des synthétases mitochondriales humaines, peu de choses sont connues. Parmi la vingtaine de synthétases dont les séquences sont maintenant identifiées moins d'une dizaine sont étudiées. Les résultats obtenus jusqu'à présent font état d'un relâchement de la spécificité de la réaction d'aminocyclation dans le contexte mitochondrial. Ce constat ne permet pas de dire s'il s'agit d'une propriété intrinsèque des systèmes d'aminocyclation mitochondriaux, résultat du confinement de ce système, du faible nombre d'ARNt mitochondriaux présents ou de leur évolution rapide par rapport aux synthétases mitochondriales codées par le génome nucléaire. Quelle qu'en soit la cause, la question se pose de la fidélité de la traduction mitochondriale. Quel est l'effet d'erreurs éventuelles de la traduction sur les fonctions assurées par la mitochondrie ? Dans une telle situation, existe-t-il un mécanisme de prise en charge des aminoacyl-ARNt erronés. Il est aussi envisageable que la spécificité d'aminocyclation soit complétée par des éléments différents des éléments canoniques connus pour les systèmes d'aminocyclation classiques. Un autre point intrigant des systèmes d'aminocyclation mitochondriaux est leur faible efficacité comparée aux systèmes cytosoliques ou eubactériens. Ceci pourrait s'expliquer par le très faible nombre de protéines mitochondriales synthétisées par la mitochondrie elle-même. D'un autre point de vue, ceci pourrait aussi être le reflet de l'existence d'un (co-)facteur essentiel pour l'aminocyclation dans la mitochondrie qui fait défaut dans les tests *in vitro*. La présence de nombreuses zones chargées positivement sur la surface de la TyrRS mitochondriale pourrait constituer un indice en faveur de cette hypothèse. Enfin, on peut aussi envisager, à l'image de la situation retrouvée dans le cytoplasme, que les synthétases mitochondriales s'associent entre elles pour améliorer leur efficacité.



# Matériel et Méthodes



# A. Matériel

## 1. Produits chimiques

$[\gamma^{32}\text{P}]\text{-ATP}$ / $[\alpha^{32}\text{P}]\text{-ATP}$ / $[\gamma^{32}\text{P}]\text{-pCp}$ / $[\alpha^{35}\text{S}]\text{-dATP}$	Amersham Pharmacia
Albumine de Sérum Bovin (BSA)	Sigma-Aldrich
Ampicilline	Carl Roth
$\beta\text{-ME}$	Carl Roth
Bleu de Coomassie R250	Carl Roth
BET	ICN
Chloramphénicol	Amersham Pharmacia
DTE	Sigma-Aldrich
DTT	Sigma-Aldrich
DTNB	Sigma-Aldrich
EDTA	Boehringer Mannheim
Glycine	Sigma-Aldrich
Imidazole	Sigma-Aldrich
IPTG	Euromedex
Kanamycine	Sigma-Aldrich
$[\text{P}^{32}]\text{-PPi}$	Perkin Elmer
Rotiphorèse Gel 30/40	Carl Roth
TBE	Euromedex
TCA	Carl Roth
TEMED	Carl Roth
Tris-base	Sigma-Aldrich
Triton X-100	Carl Roth
Tyrosine	Merck
$[\text{H}^3]$ / $[\text{C}^{14}]\text{-tyrosine}$	Amersham Pharmacia
X-Gal	Euromedex

## 2. Enzymes commerciales

<i>Afl</i> I, <i>Bam</i> HI, <i>Bgl</i> II, <i>Bst</i> MI, <i>Dpn</i> I, <i>Xmn</i> I	New England Biolabs
BAP (Bacterial Alkaline Bovine)	Fermentas
Dynazyme EXT polymérase	Finnzymes
Pfu Ultra ADN polymérase	Stratagene
SAP (Shrimp Alkaline Phosphatase)	Fermentas
Sequenase	usb
T4 ADN ligase	New England Biolabs
T4 PNK (PolyNucleotide Kinase)	Fermentas

## 3. Enzymes non commerciales

**T7 ARN polymérase** : cette enzyme permet de transcrire des gènes sous contrôle d'un promoteur spécifique. Elle est couramment utilisée pour la synthèse *in vitro* d'ARN et dans notre cas particulier d'ARNt. La T7 ARN polymérase ADN dépendante a été purifiée au laboratoire selon le protocole mis au point par (Becker et al., 1996) à partir d'un clone donné par Studier (Studier et Moffatt, 1986).

**TyrRS cytoplasmique humaine** : le vecteur d'expression contenant cette enzyme nous a été généreusement fournie par P. Schimmel (Scripps Institute, USA, CA, La Jolla).

## 4. Souches d'*Escherichia coli*

**TG1** : *F'* [*traD36 proAB<sup>+</sup> lacI<sup>f</sup> lacZΔM15*] *supE Δ(hsdM-mcrB)5 thi Δ(lac-proAB)*. Cette souche a servi lors du clonage des ARNt.

**TOP10** : *F mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 deoR recA1 araD139, Δ(ara-leu)7697, galU, galk, rpsL, endA1, nupG*. Cette souche a été utilisée lors du clonage des ARNt et de la mt-TyrRS humaine.

**TOP10F'** : *F'* [*lacI<sup>f</sup> Tn10 (Tet<sup>r</sup>)*] *mcrA Δ(mrr-hsdRMS-mcrBC) φlacZΔM15 ΔlacX74 recA1 araD139 Δ(ara-leu)7697 galU galk rpsL (Str<sup>R</sup>) endA1 nupG*. Cette souche est fournie compétente dans le kit "TA-cloning" (Invitrogen) et a été transformée avec les clonages de produit de PCR dans le vecteur pCR<sup>®</sup>2.1.



**BL21-codon plus RILX** : *F<sup>-</sup>, ompT, hsdS(sB<sup>-</sup> mB<sup>-</sup>), dcm<sup>+</sup>, Tetr, gal, endA, Hte, [argU, ileY, leuW, Camr] metA:Tn5(Kan<sup>r</sup>)*. Cette souche est transformée par un plasmide qui code les ARNt capables de s'apparier aux codons rares spécifiant l'Arg, l'Ile et la Leu, ce qui permet d'améliorer l'expression du gène recombiné. Parallèlement elle est rendue auxotrophe pour la Met par l'interruption du gène codant l'homosérine o-succinyltransférase par le transposon Tn5. Cette souche est utilisée pour surproduire des protéines dont les résidus Met sont substitués par la SeMet (sélénométhionine). Les protéines sélénométhionylées sont utilisées en cristallographie pour résoudre leur structure tridimensionnelle par la technique MAD (Multiwavelength Anomalous Diffraction).

## 5. Plasmides

**pCR<sup>®</sup>2.1** : contient une origine de réplication, les gènes de résistance à l'ampicilline (Amp<sup>R</sup>) et à la kanamycine (Kan<sup>R</sup>), le promoteur *lac* et le gène *lacZ $\alpha$* . Il est fourni linéaire au niveau du gène *lacZ $\alpha$*  et présente des T protrudants aux extrémités 5' et 3' (kit "TA Cloning" Invitrogen). Il permet le clonage des produits de PCR qui possèdent à leurs extrémités 3' des A sortants. L'intérêt de ce vecteur est d'identifier les clones recombinants par le système de sélection blanc/bleu. Ce système est basé sur l'interruption du gène *lacZ $\alpha$*  codant pour la  $\beta$ -galactosidase par l'insertion du produit de PCR.

**pQE30** : permet l'expression dans *E. coli* de protéines avec une séquence poly-His en N-terminal, qui permettra leur rétention sur colonne Ni-NTA. Le vecteur possède le gène de résistance à l'ampicilline (Amp<sup>R</sup>).

**pQE70** : permet l'expression dans *E. coli* de protéines avec une séquence poly-His en C-terminal, qui permettra leur rétention sur colonne Ni-NTA. Le vecteur possède le gène de résistance à l'ampicilline (Amp<sup>R</sup>).

**pUC119** : dérivé du vecteur pUC18 (Yanisch-Perron et al., 1985) comprenant une origine de réplication, le gène de résistance à l'ampicilline (Amp<sup>R</sup>).

## 6. Milieux de culture

Les souches d'*E. coli* sont cultivées à 37°C dans le milieu LB (Luria Bertani) contenant pour 1 litre 10 g de bactotryptone, 5 g d'extraits de levure et 10 g de NaCl. Les milieux gélosés sont complétés par 20 g de bactoagar. Le milieu est additionné, si nécessaire et après autoclavage, de 100 µg/mL d'ampicilline, 170 µg/mL de chloramphénicol, ou de 50 µg/mL de kanamycine.

Pour la production de protéines sélénométhionylées, la souche RILX est cultivée dans du milieu M9 (contenant pour 1 litre 8 g de Na<sub>2</sub>HPO<sub>4</sub>, 4 g de KH<sub>2</sub>PO<sub>4</sub>, 0,5 g de NaCl et 0,5 g de NH<sub>4</sub>Cl) additionné de 0,4% de glucose, 1 mM de MgSO<sub>4</sub> et 0,3 mM de CaCl<sub>2</sub>. Après autoclavage, le milieu est encore supplémenté avec 1 mg/L de biotine, 1 mg/L de thiamine, un mélange de 7 acides aminés essentiels à 100 mg/L (Phe, Leu, Ile, Lys, Val, Thr et Trp) et une solution d'oligoéléments diluée 1000x (contenant pour un litre 50 g d'EDTA ajusté à pH 7,5, 840 mg de ZnCl<sub>2</sub>, 8,3 g de FeCl<sub>3</sub>, 130 mg de CuCl<sub>2</sub>, 100 mg de CoCl<sub>2</sub>, 100 mg de H<sub>3</sub>BO<sub>3</sub> et 16 mg de MnCl<sub>2</sub>). Cette préparation est dénommée milieu M9 "modifié". Enfin les antibiotiques sont ajoutés aux concentrations finales de 100 µg/mL pour l'ampicilline et 170 µg/mL pour le chloramphénicol.

## **B. Techniques courantes**

### **1. Préparation et transformation de bactéries compétentes**

#### **1.1. Méthode au CaCl<sub>2</sub>**

Cette technique consiste à exposer les cellules en phase exponentielle de croissance au CaCl<sub>2</sub> qui perméabilise les membranes et facilite ainsi la pénétration d'ADN exogène dans les bactéries. Une colonie de bactéries TG1 ayant poussé sur du milieu LB gélosé est prélevée stérilement et mise en culture dans 25 mL de milieu LB à 37°C, sous agitation, jusqu'à atteindre la phase exponentielle de croissance ( $DO_{600}=0,5$ ). Après centrifugation pendant 10 min à 4 000 *g* et 4°C le culot de cellules est repris dans 15 mL de CaCl<sub>2</sub> 50 mM froid et incubé 30 min dans la glace. Les cellules sont à nouveau centrifugées pendant 15 min à 2 000 *g* et 4°C et culot repris dans 2 mL de CaCl<sub>2</sub> 50 mM. Après ajout de 1,2 mL de glycérol 100%, les cellules sont aliquotées par volume de 200 µL et directement utilisées ou stockées à -80°C pour un usage ultérieur.

La transformation de bactéries ainsi rendues compétentes s'effectue par l'ajout de l'ADN à intégrer. Le milieu est incubé 1 h à 4°C puis les bactéries sont soumises à un choc thermique de 1 min à 42°C et remises dans la glace à 4°C pendant 5 min. Le mélange est alors étalé sur boîte de Pétri contenant du milieu LB gélosé supplémenté avec le(s) antibiotique(s) adéquat(s). Les différents temps d'incubations peuvent varier en fonction de la souche bactérienne utilisée.

#### **1.2. Electrocompétence**

Une colonie de bactéries TG1 ou Top10 ayant poussé sur du milieu LB gélosé est prélevée stérilement et mise en pré-culture dans 5 mL de milieu LB à 37°C pendant une nuit. Cette pré-culture est utilisée pour ensemercer 1 L de milieu LB, qui sera mis sous agitation à 37°C jusqu'à atteindre la phase exponentielle de croissance ( $DO_{600}=0,5$ ). Les cellules sont centrifugées pendant 15 min à 4 000 *g* et 4°C, puis lavées 2 fois avec 500 mL de HEPES-NaOH 1 mM froid. Les culots sont alors rassemblés et lavés dans 20 mL de glycérol 10% froid, et repris dans un volume final de 1,2 mL de glycérol 10% froid. Les cellules sont conservées à -80°C par aliquotes de 50 µL.

La transformation de bactéries ainsi préparées se fait par addition d'ADN aliquoté dans un milieu débarrassé de toute trace de sels. Le mélange est transféré dans une cuve à électroporation et soumis à une décharge électrique (2 500 V, 200  $\Omega$  et 25  $\mu$ F) qui va perturber la membrane des bactéries et ainsi faciliter la pénétration de l'ADN exogène. Cette opération est suivie de l'ajout immédiat de 300  $\mu$ L de milieu LB. Différentes dilutions sont étalées sur milieu LB gélosé contenant le(s) antibiotique(s) adéquat(s).

## **2. Étude des acides nucléiques**

### **2.1. Détermination de la concentration en acides nucléiques**

Les acides nucléiques absorbent la lumière UV et présentent un pic d'absorption à 260 nm. La concentration d'une solution d'ADN est estimée par la relation : concentration (mg/mL) = 0,05 x  $DO_{260}$ , tandis que la concentration en ARN est estimée par la relation : concentration (mg/mL) = 0,04 x  $DO_{260}$ . Sur un spectre d'absorption réalisé entre 220 et 320 nm, une solution pure d'acide nucléique présente un maximum d'absorption à 256-260 nm et un rapport  $A_{260}/A_{280}$  compris entre 1,8 et 2.

### **2.2. Extraction phénolique**

Cette étape permet de séparer les acides nucléiques des protéines. L'échantillon à traiter est additionné d'un volume de phénol, vortexé pendant 1 min et centrifugé 5 min à 10 000 *g*. La phase aqueuse est soigneusement prélevée, de façon à ne pas entraîner les protéines dénaturées à l'interface avec la phase phénolique, lavée par addition de 2 volumes d'éther saturé en eau, agitée vigoureusement pendant 1 min et centrifugée 1 min à 10 000 *g*. Après élimination de la phase étherée, les traces d'éther résiduelles sont évaporées. Les acides nucléiques contenus dans la phase aqueuse sont généralement précipités.

### **2.3. Précipitation des acides nucléiques**

La solution d'acides nucléiques à précipiter est additionnée de 0,1 volume d'acétate de sodium 3 M (pH 5,0 à 6,0) et de 2,5 volumes d'éthanol absolu 100% puis placée à -20°C

pendant 20 min à 1 h à -80°C. Après 30 min de centrifugation à 12 000 *g* à 4°C, le culot est lavé avec de l'éthanol 70% et centrifugé une seconde fois. Le culot est séché sous vide et repris dans de l'eau.

## **2.4. Amplification d'ADN plasmidique**

L'ADN plasmidique est amplifié par culture de la souche bactérienne choisie dans 100 mL de milieu LB en présence de l'antibiotique adéquat. La purification de l'ADN plasmidique est réalisée à l'aide d'un kit commercialisé (Plasmid MAXI kit, Genomed), suivant le protocole fourni. Brièvement, après lyse des cellules, et élimination des débris par centrifugation, l'ADN plasmidique (présent dans le surnageant) est retenu sur colonne échangeuse d'anions. L'ADN est ensuite élué et précipité par ajout de 0,7 volumes d'isopropanol.

## **2.5. Migration des acides nucléiques sur gels**

### ***Electrophorèse sur gel d'agarose***

Le gel d'agarose contient habituellement 1 % d'agarose (p/v), dissout à chaud dans du tampon TBE 1x (0,1 M Tris-base pH 8,3, 0,1 M acide borique, 2 mM EDTA). Les échantillons d'ADN, après addition d'un volume de tampon de dépôt (80 % (v/v) formamide, 10 mM EDTA pH 8,0, 0,025 % (p/v) xylène cyanol et bleu de bromophénol) sont déposés sur le gel et séparés par migration à 100 V dans une solution de TBE 1x. Les acides nucléiques sont visualisés sous UV après un bain dans une solution de BET à 0,5 µg/mL, TBE 1x.

### ***Electrophorèse sur gel de polyacrylamide***

L'électrophorèse sur gel de polyacrylamide est une méthode largement utilisée pour la purification et l'analyse d'acides nucléiques. Le gel de polyacrylamide est un copolymère constitué d'acrylamide et de N-N'-méthylène bis-acrylamide dont le rapport est de 19/1 (Rotiphorèse Gel 40). Le gel est constitué de 6 à 22 % d'acrylamide/bis-acrylamide et de TBE 1x. Dans le cas des gels dénaturants, de l'urée est ajoutée à une concentration finale de 8 M. La polymérisation du gel est déclenchée par l'ajout de 0,5 % (v/v) de solution de peroxydisulfate d'ammonium à 10 % et de 0,1 % (v/v) de TEMED. L'épaisseur du gel varie de 0,4 mm pour les gels analytiques ou de séquençage à 3 mm pour les gels préparatifs (utilisés lors de la purification de grande quantité d'ARN). Avant dépôt, les échantillons sont additionnés d'un volume de tampon de dépôt (20 % (p/v) saccharose, 8 M urée, 0,025 % (p/v) xylène cyanol et bleu de bromophénol). L'électrophorèse est réalisée dans du tampon

TBE 1x. La migration des échantillons s'effectue à 150 V sur mini-gel (0,075x8,5x10,5 cm), à 1 800 V et 55 mA sur gel de séquence ou gel de purification (0,075x30x40 cm), et à 650 V et 55 mA sur gels préparatifs (0,2x30x40). La durée de migration varie en fonction de la taille de l'ADN ou de l'ARN étudié.

## **2.6. Éluion des ARN à partir d'un gel de polyacrylamide**

Après migration sur gel de polyacrylamide, les ARN sont visualisés par ombrage aux UV (à 254 ou 312 nm) sur un écran de fluorographie. Les transcrits radiomarqués sont visualisés par autoradiographie. La bande de gel contenant la molécule de taille attendue est découpée et les ARN sont extraits du gel de manière passive ou par électroélution.

### ***Elution passive***

Le fragment de gel contenant l'ARN est incubé 4 h à 4°C dans 500 µL de tampon d'éluion (0,5 M acétate d'ammonium pH 6,1, 10 mM acétate de magnésium, 0,1 mM Na<sub>2</sub>EDTA et 0,1 % SDS), sous agitation. La phase aqueuse est récupérée et l'ARN est précipité.

### ***Electroélution***

Cette méthode est utilisée lorsque l'épaisseur du gel utilisé pour l'électrophorèse est trop importante pour permettre une éluion passive efficace. Les bandes de gel contenant l'ARN sont placées dans un appareil à électroélution "Biotrap" (Schleicher & Schuell) muni de membranes semi-perméables, et immergées dans du tampon TBE 1x stérile. En appliquant un champ électrique de 150 V pendant 1 h 30 à 4°C, les ARN sont extraits du gel et récupérés dans la chambre anodique. Cette opération est renouvelée et permet de récupérer 90 à 100 % du matériel biologique. Les éluats sont ensuite récupérés et précipités pour être concentrés.

## **3. Étude des protéines**

### **3.1. Détermination de la concentration en protéines**

Les protéines peuvent être quantifiées par dosage colorimétrique selon la méthode de Bradford (réactif Bio-Rad) en utilisant l'albumine de sérum bovin (BSA) comme protéine étalon. La concentration en protéine pure peut également être calculée en fonction du coefficient d'extinction molaire de celle-ci, selon la relation de Beer-Lambert :  $c \text{ (mg/mL)} =$

$DO_{280} \times E^{-1} \times I^{-1}$ , où E est le coefficient d'extinction molaire ( $M^{-1}.cm^{-1}$ ) et I la distance optique. Les coefficients d'extinction molaire pour les protéines clonées sont de 41 500  $M^{-1}.cm^{-1}$  pour la mt-TyrRS, 39 250  $M^{-1}.cm^{-1}$  pour la mt-TyrRS- $\Delta$ S4 et 8 480  $M^{-1}.cm^{-1}$  pour la mt-TyrRS-S4.

### **3.2. Électrophorèse en conditions dénaturantes**

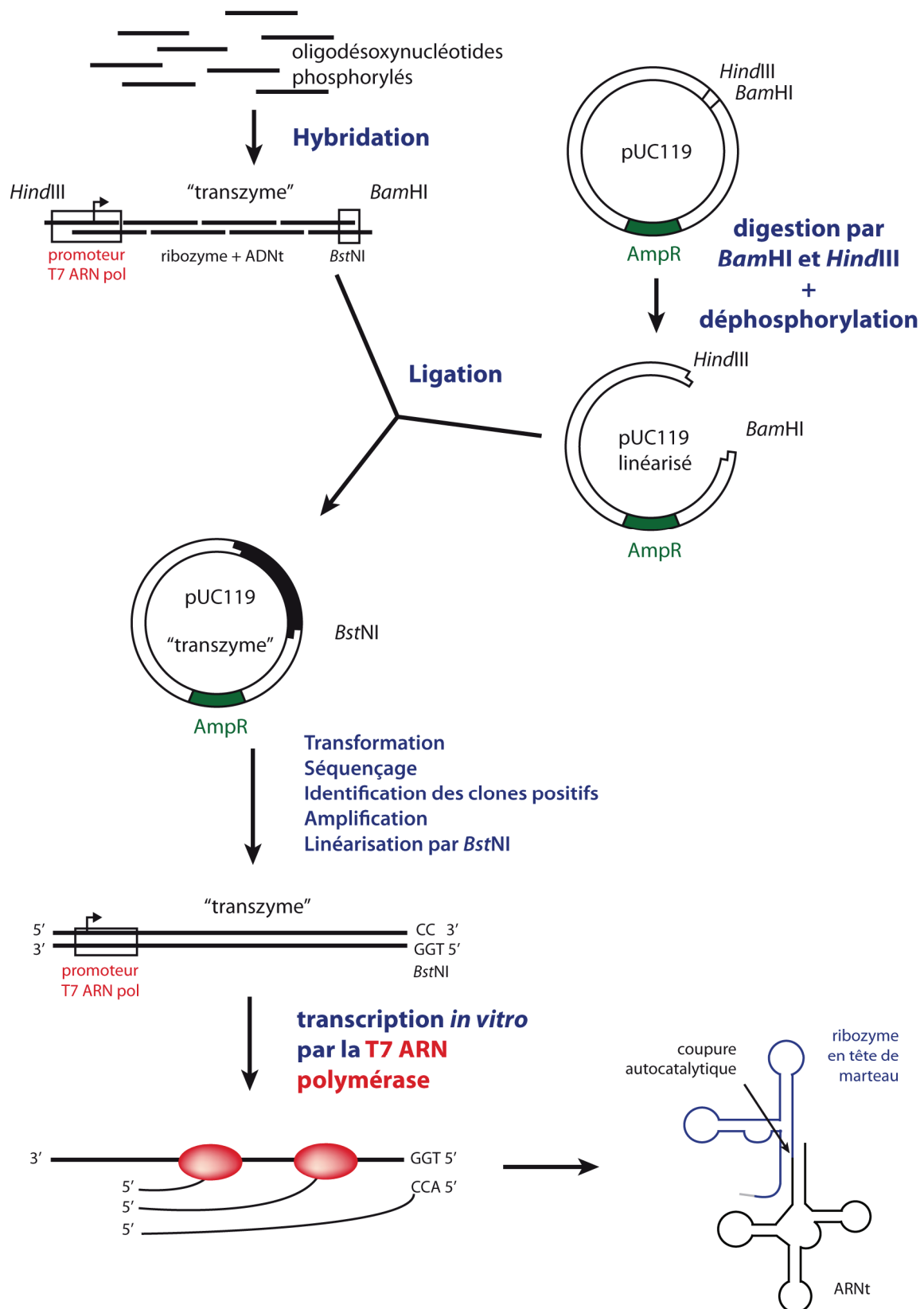
Cette technique permet d'évaluer la pureté et le poids moléculaire apparent de protéines en solution. Le gel de polyacrylamide est constitué (i) d'un gel de concentration à faible maillage qui permet la concentration des protéines de l'échantillon et (ii) d'un gel de résolution qui permet de séparer les protéines en fonction de leur poids moléculaire apparent. Les deux parties du gel sont constituées d'acrylamide et de N-N'-méthylène bis-acrylamide suivant un rapport de 37,5/1 (Rotiphorèse Gel 30). Le gel de concentration est constitué de 5 % d'acrylamide/bis-acrylamide, 0,1 % de SDS et de 125 mM de tampon Tris-HCl pH 6,8. Le gel de résolution est constitué de 6 à 15 % d'acrylamide/bis-acrylamide, 0,1 % de SDS et de 375 mM de tampon Tris-HCl pH 8,8. La migration des protéines s'effectue dans un tampon 25 mM de Tris-base, 140 mM de glycine et 0,1 % SDS. Avant dépôt sur le gel, les échantillons sont additionnés d'un volume de solution de dépôt (50 mM Tris-HCl pH 6,8, 2 % SDS, 0,1 % bleu de bromophénol, 20 % glycérol) et chauffés 5 min à 100°C. Le gel peut être fait en conditions réductrices par l'ajout de 100 mM de DTE au tampon de dépôt. Le voltage appliqué pour la migration est d'abord de 80 V pendant la migration des protéines à travers la couche de concentration puis de 150 V pour la couche de résolution. Les protéines présentes dans le gel sont visualisées par traitement du gel avec une solution de bleu de Coomassie (0,25 % (p/v) bleu de Coomassie R250, 45 % méthanol, 10 % acide acétique) pendant 15 min. La décoloration du gel se fait par lavages successifs dans une solution contenant 45 % d'éthanol et 10 % d'acide acétique jusqu'à visualisation des bandes de protéines.

### **3.4. Diffusion de lumière**

L'analyse par diffusion de lumière apporte des renseignements sur la taille apparente des particules macromoléculaires en solution et leur homogénéité (polydispersité et agrégation). L'emploi de cette méthode est devenue un prérequis réalisé de manière classique pour la cristallogénèse des macromolécules biologiques. Les mesures de coefficient de diffusion sont effectuées à l'aide d'un appareil DynaPro (Wyatt Technology, Santa Barbara, CA). L'échantillon protéique (12  $\mu$ L à une concentration comprise entre 1 et 4

mg/mL) est injecté dans une cellule en quartz, centrifugé 10 min à 5 000 *g* pour supprimer les agrégats de grande taille, puis irradié par un faisceau laser ( $\lambda=780$  nm). Une fonction d'autocorrélation exponentielle décroissante est calculée à partir du signal de diffusion. Le temps de décroissance est proportionnel au coefficient de diffusion de translation  $D_T$  des particules. En assimilant les particules à des sphères dures, la relation de Stokes-Einstein  $R_h = kT/6\pi\eta D_T$  (avec  $k$  la constante de Boltzmann,  $T$  la température et  $\eta$  la viscosité) permet de calculer leur rayon hydrodynamique  $R_h$ . Le programme Dynamics indique les valeurs de  $D_T$  et  $R_h$  ainsi que l'écart type sur le rayon  $R_h$  servant de mesure à la polydispersité. En général une protéine pure et homogène utilisable en cristallogénèse est monodisperse, alors que les protéines agrégées ou contaminées sont polydisperses.





**Figure 26. Production de "transzymes"**

Principe de clonage des "transzymes" et de transcription *in vitro* par la T7 ARN polymérase.

## C. Préparation d'ARNt par clonage et transcription *in vitro*

### 1. Principe

Les gènes synthétiques des différents ARNt<sup>Tyr</sup> que nous avons choisis d'étudier sont clonés sous forme de "transzyme" dans le vecteur pUC119 (Fechter et al., 1998). Cette technique, résumée dans la **figure 26**, permet de transcrire de manière efficace un gène dont les premiers nucléotides de la séquence sont défavorables à la transcription par la T7 ARN polymérase (Dunn et Studier, 1983). Les "transzymes" (pour **transfert** et **ribozyme**) comportent successivement (i) la séquence promotrice optimale reconnue par la T7 ARN polymérase, (ii) la séquence d'un ribozyme en tête de marteau (Price et al., 1995), (iii) la séquence de l'ARNt concerné et (iv) le site de restriction *Bsa*NI. Ce dernier permet l'obtention, après transcription, de molécules se terminant en 3' par la séquence -CCA<sub>OH</sub> indispensable à l'aminacylation. L'activité autocatalytique du "transzyme" libère deux molécules d'ARN : le ribozyme (50 nts) et l'ARNt (76 nts).

### 2. Construction des gènes synthétiques

Les gènes "transzymes" des séquences sauvage et mutantes de l'ARNt<sup>Tyr</sup> mitochondrial humain sont construits à partir de 8 oligonucléotides de synthèse chevauchants. Ils doivent être phosphorylés à leur extrémité 5' pour permettre leur hybridation. La phosphorylation se fait à 37°C pendant 15 min dans un milieu réactionnel de 50 µL contenant 50 mM Tris-HCl pH 7,6, 10 mM MgCl<sub>2</sub>, 5 mM DTE, 1 mM ATP, 400 pmoles de chaque oligonucléotide et 10 U de T4 polynucléotide kinase. L'enzyme est ensuite inactivée par une incubation de 10 min à 65°C. L'hybridation des oligonucléotides phosphorylés (16 pmoles de chaque oligonucléotide) se fait dans un volume de 40 µL contenant 50 mM Tris-HCl pH 7,6, 10 mM MgCl<sub>2</sub>, 5 mM DTE. Le milieu réactionnel est chauffé à 80°C pendant 10 min puis subit un refroidissement lent jusqu'à 25°C (2 à 3 h).

### 3. Clonage

Les gènes synthétiques des "transzymes" sont clonés dans le vecteur pUC119 entre les sites de restriction *Bam*HI et *Hind*III. Ce vecteur confère à la souche hôte la résistance à l'ampicilline. La digestion du vecteur (10 µg) se fait pendant 3 h à 37°C dans 100 µL, en présence de 40 U de *Bam*HI, 40 U de *Hind*III, 10 mM Tris-HCl pH 7,9, 10 mM MgCl<sub>2</sub>, 50 mM NaCl, et 1 mM DTE. Le vecteur digéré est récupéré par extraction phénolique puis précipité à l'éthanol. Le culot est lavé, séché et repris dans 50 µL de tampon 10 mM Tris-HCl pH 7,5, 10 mM MgCl<sub>2</sub>. L'ajout de 600 U de BAP pendant 30 min à 60°C permet de déphosphoryler le vecteur, afin de prévenir la recircularisation du vecteur sur lui-même. Celui-ci est récupéré par extraction phénolique, précipité à l'éthanol et resolubilisé dans 20 µL d'eau stérile.

La ligation des oligonucléotides au vecteur est effectuée à 16°C pendant 16 h en incubant 0,1 µg de plasmide pUC119 linéarisé et déphosphorylé avec 4 pmoles de "transzyme" (oligonucléotides hybridés) prélevés dans le tube d'hybridation, en présence de 5 U de T4 ADN ligase et 2,5 mM d'ATP dans 20 µL.

La transformation de 200 µL de bactéries TG1 rendues compétentes au CaCl<sub>2</sub> par 15 µL de milieu de ligation se fait pendant 5 min dans la glace. Les bactéries sont ensuite soumises à un choc thermique de 30 sec à 42°C, et remises 2 min dans la glace. Le milieu de transformation est étalé sur une boîte de Pétri contenant du milieu gélosé LB/Amp, et la boîte est placée dans une étuve à 37°C pendant 16 h.

### 4. Sélection des clones positifs

La recherche de clones positifs se fait par séquençage des plasmides. Une dizaine de colonies est prélevée sur la boîte et mise en culture dans 3 mL de milieu LB/Amp pendant 5 h. L'ADN plasmidique (environ 10 µg), obtenu par mini-préparation (kit commercialisé QIAprep Spin, QIAGEN), est dénaturé en présence de 0,2 M NaOH pendant 10 min à 25°C, puis précipité à l'éthanol. Le séquençage se fait selon le protocole commercialisé de la Séquenase® 2.0 (United States Biochemicals), en hybridant une amorce spécifique du vecteur pUC119 en amont du gène inséré ("Forward" 5'-GTA AAA CGA CGG CCA GT-3' ou "Reverse" 5'-GGA AAC AGC TAT GAC CAT G-3' situées de part et d'autre de la cassette de clonage). Les produits de séquençage radiomarqués par l'incorporation de [ $\alpha$ -<sup>35</sup>S]dATP sont séparés par migration (1 800 V) sur un gel 8% acrylamide/bisacrylamide (19/1, Rotiphorese

Gel 40), 8 M urée, et TBE 1x. Le gel est séché et les séquences sont révélées par autoradiographie.

Les clones positifs sont amplifiés par culture de la souche bactérienne correspondante dans 100 mL de milieu LB/Amp. La purification à grande échelle de l'ADN plasmidique est réalisée à l'aide d'un kit commercial (Jet Star, Genomed) comme décrit au paragraphe **B.2.4**.

## **5. Transcription *in vitro***

La transcription de 50 µg d'ADN matrice par 2,5 µg de T7 ARN polymérase se fait pendant 3h30 à 37°C dans un milieu réactionnel de 500 µL composé de 40 mM Tris-HCl pH 8,1, 22 mM MgCl<sub>2</sub>, 0,01% Triton X-100, 1 mM spermidine, 5 mM DTE, et 4 mM de chacun des 4 NTPs. La coupure autocatalytique du "transzyme" n'est pas optimisée comme décrit (Fechter et al., 1998), les ARNt mitochondriaux étant partiellement dégradés lors du chauffage à 65°C. La réaction est arrêtée par une extraction phénolique suivie d'une précipitation à l'éthanol.

## **6. Purification des transcrits sur gels préparatifs**

Le transcrit correspondant à l'ARNt est contaminé non seulement par l'ADN matrice, mais aussi par les nucléotides non incorporés, le pyrophosphate libéré, les transcrits incomplets, le "transzyme" non clivé (140 nucléotides) et le ribozyme libéré après clivage (50 nucléotides). La purification du transcrit de l'ARNt se fait en fonction de la taille par migration du produit de transcription (repris dans 200 µL de tampon de dépôt) sur un gel préparatif 12% acrylamide/bisacrylamide (19/1, Rotiphorese Gel 40), 8 M urée, et TBE 1x (650 V, 65 mA, 60 W pendant 16 h). La bande correspondant au transcrit d'ARNt est détectée par ombrage aux UV et découpée. L'ARN est extrait du gel par électroélution comme décrit au paragraphe **B.2.6**.

## 7. Analyse de l'extrémité 5' des ARNt transcrits

### 7.1. Principe

La T7 ARN polymérase introduit *in vitro* des erreurs en fin de transcription (Milligan et al., 1987). La présence de molécules incorrectes en 3' fausse les tests fonctionnels puisque seuls les ARNt présentant une extrémité 3' -CCA sont aminoacylables. Afin de déterminer la proportion de molécules actives, la nature du nucléotide terminal des ARNt transcrits est déterminée.

### 7.2. Marquage de l'extrémité 3' des transcrits au [ $\gamma$ -<sup>32</sup>P]-pCp

Le [ $\gamma$ -<sup>32</sup>P]-pCp est fixé en 3' du transcrit grâce à l'action de la T4 ARN ligase. La réaction se fait sur 50 pmoles d'ARNt dans 20  $\mu$ L de milieu réactionnel comprenant le tampon T4 ARN ligase (10 mM HEPES-NaOH pH 7,5, 20 mM MgCl<sub>2</sub>, 10 mM DTE, 12% DMSO, 100  $\mu$ M ATP), et 10  $\mu$ Ci de [ $\gamma$ -<sup>32</sup>P]-pCp, pendant 16 h à 4°C. Les produits de la réaction sont déposés sur un gel dénaturant (2 mm) 12% acrylamide/bisacrylamide (19/1, Rotiphorese Gel 40), 8 M urée et TBE 1x. La migration se fait pendant 1h30 (1 500 V, 55 mA, 60 W). Après autoradiographie, la bande de gel renfermant le transcrit est découpée et l'ARN est extrait par une élution passive, comme décrit au paragraphe **B.2.6**. L'éluat est précipité à l'éthanol et la radioactivité du culot mesurée par un comptage Cerenkov. Le culot est repris dans un volume d'eau stérile afin d'obtenir 30 000 cpm/ $\mu$ L.

### 7.3. Hydrolyse des transcrits

Une aliquote (1  $\mu$ L) de l'éluat obtenu, additionné de 10  $\mu$ g d'ARNt total de levure, est soumis à une hydrolyse en présence de pipéridine 15% dans un volume réactionnel de 20  $\mu$ L. La réaction est réalisée pendant 2 h à 95°C dans un capillaire scellé, afin que l'hydrolyse soit totale et libère des nucléosides 5' phosphate.

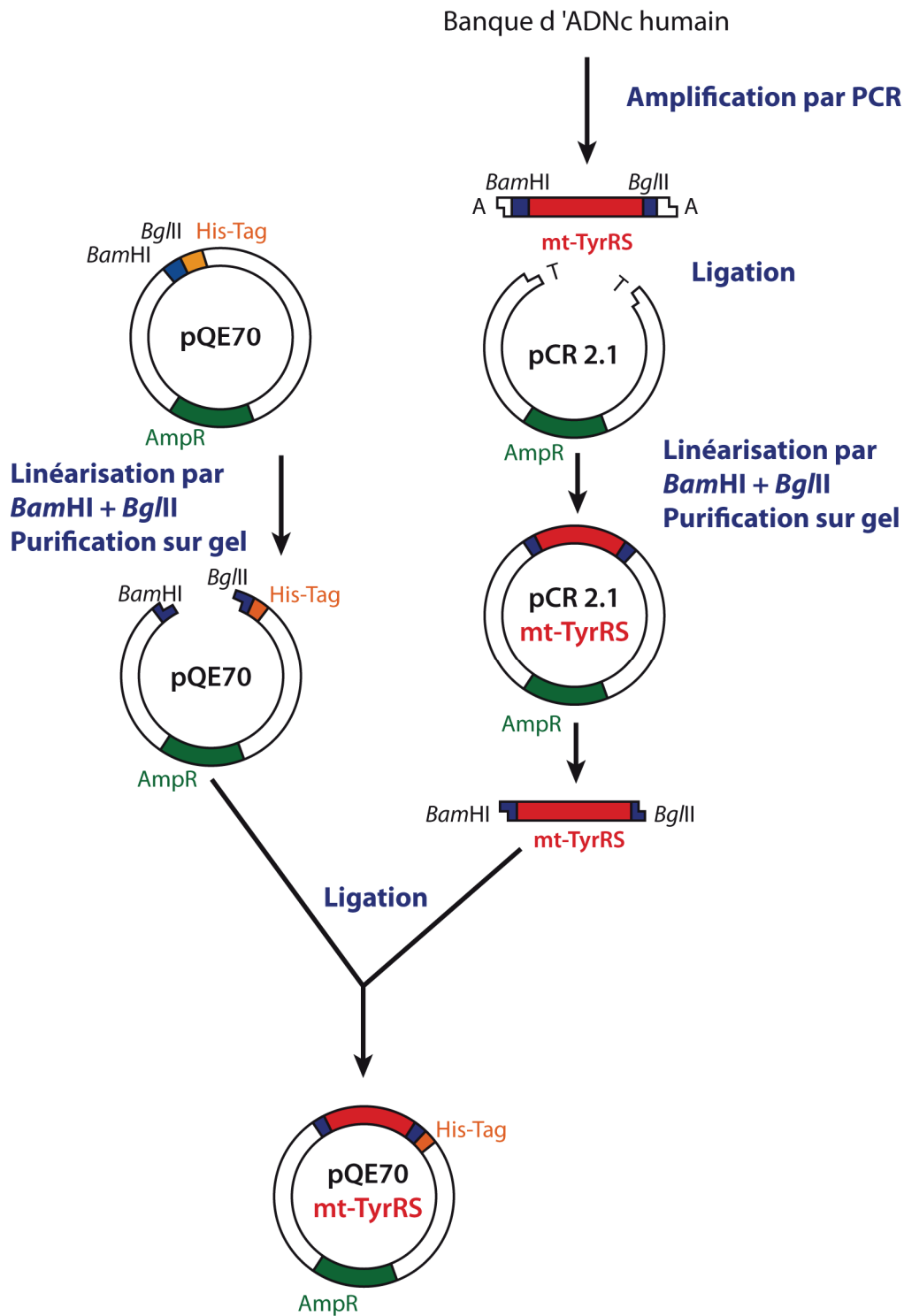
### 7.4. Chromatographie bidimensionnelle

L'hydrolysate est déposé avec précaution  $\mu$ L par  $\mu$ L dans un coin d'une plaque chromatographique de 10x10 cm<sup>2</sup> (Ready Foils). Une chromatographie à 2 dimensions est

nécessaire pour différencier les 4 nucléotides. La première migration se fait pendant 3 h avec un solvant "isobutyrique" (66% acide isobutyrique, 0,25% ammoniacque, 1 mM EDTA). Après séchage de la plaque, la deuxième migration se fait perpendiculairement à la première dans un solvant "isopropanol" (68% isopropanol, 6,5% HCl, 1 mM EDTA). Après 4 h de migration, la plaque est à nouveau séchée.

## **7.5. Révélation**

Une autoradiographie par BioImager® (Genomic Solutions) après impression d'une "Image Plate" (Fuji Photo Film, Inc.) pendant 1 h permet de détecter le (ou les) nucléotide(s) terminal(ux) de chaque lot de transcrit. Le traitement informatique des données (MacBAS®) permet de quantifier chaque tache de la plaque et de calculer la proportion de transcrit se terminant par le nucléotide attendue.



**Figure 27. Principe du clonage de la mt-TyrRS humaine dans le vecteur d'expression pQE70**

Ce plasmide permet d'ajouter une séquence de 6 résidus histidine (His-Tag) à l'extrémité C-terminale de la séquence clonée.

## **D. Clonage, surexpression et purification de la TyrRS mitochondriale humaine et de ses variants**

### **1. Clonage du gène à partir d'une banque d'ADNc humains**

#### **1.1. Amplification du gène par PCR**

Le gène de la TyrRS mitochondriale humaine a été amplifié par PCR à partir d'une banque d'ADNc humain (F. Martin, IBMC). L'ajout préalable des sites de restriction *Bam*HI et *Bgl*II aux oligonucléotides servant d'amorces permet l'insertion du gène dans le vecteur pQE70 (5'-GGA TCC ACA CTC GGG CGC TCA GGG GTT ACT GG-3' en 5' avec le site *Bam*HI et 5'-AGA TCT CAA CTG AAG CCA TTT TAT AAT G-3' en 3' avec le site *Bgl*II). Les conditions de la PCR sont de 5 min à 95°C suivi de 25 cycles de 1 min à 95°C, 1 min à 55°C et 1 min à 70°C, finalisé par 10 min à 70°C. Le milieu réactionnel pour la PCR (50 µL) contient 2 µg d'ADNc issu de la banque (matrice), 1 µM d'oligonucléotides, 200 nM de chaque dNTP, 1,5 mM de MgCl<sub>2</sub>, 10 µL de tampon EXT (50 mM Tris-HCl pH 9, 15 mM sulfate d'ammonium, 0,1% Triton X-100) et 2 µL de *Taq* polymérase (Dynazine EXT).

#### **1.2. Clonage dans le vecteur pCR2.1**

Lors de la PCR, la *Taq* polymérase ajoute spécifiquement des résidus désoxyadénosine (dA) en 3' des brins polymérisés. Le vecteur pCR2.1, fourni linéarisé dans le kit "TA cloning" (Stratagène), présente quant à lui un résidu désoxythymidine (dT) non apparié en 3' de ses 2 brins, qui permet l'intégration des fragments de la PCR. La ligation (10 µL) se fait pendant 16 h à 16°C en incubant 1 µL d'une dilution du produit de la PCR (pour obtenir un ratio 1:1 insert/vecteur) avec 50 ng de vecteur pCR2.1 linéarisé et 4 U de T4 ADN ligase dans le tampon de ligation fourni. Une aliquote de 100 µL de bactéries compétentes TOP10F' est transformée par 2 µL du milieu de ligation additionné de 2 µL de β-ME à 0,5 M par incubation dans la glace pendant 1 h suivie d'un chauffage de 30 sec à 42°C. Les cellules sont remises en culture avec 250 µL de milieu SOC (2% tryptone, 0,5% d'extraits de levure, 10 mM NaCl, 2,5 mM KCl, 10 mM MgCl<sub>2</sub>, 10 mM MgSO<sub>4</sub>, 20 mM glucose) pendant 1 h à 37°C sous agitation. Des aliquotes de 50 et 200 µL sont finalement étalées sur boîtes de Pétri



LB/Amp additionnées de 40  $\mu$ L de X-Gal à 20 mg/mL et 40  $\mu$ L d'IPTG à 0,1 M. Les boîtes sont placées dans une étuve à 37°C pendant 16 h, puis à 4°C pendant 6 h. Les clones positifs sont repérés par leur absence de coloration bleue. La présence de l'insert est vérifiée par PCR en utilisant les amorces décrites dans le protocole du kit. L'intégrité de la séquence est vérifiée par séquençage automatique.

### **1.3. Sous-clonage dans le vecteur pQE70**

#### **1.3.1. Purification de l'insert**

L'ADN plasmidique du vecteur pCR2.1 est digéré respectivement par *Bam*HI et *Bgl*II pour le clonage dans pQE70. Le milieu réactionnel de 100  $\mu$ L contient 5  $\mu$ g d'ADN, 50 U de chaque enzyme, le tampon *Bam*HI (10 mM Tris-HCl pH 7,9, 150 mM NaCl, 10 mM MgCl<sub>2</sub>, 1 mM DTE et 100  $\mu$ g/mL de BSA) et est incubé 2 h à 37°C. La réaction est arrêtée par une extraction phénolique, puis l'ADN est précipité en présence de 0,3 M d'AcONa pH 6,0 et de 2,5 volumes d'éthanol. Le culot d'ADN est lavé, séché, repris dans 10  $\mu$ L d'eau stérile, additionné d'un volume de tampon de dépôt (50 mM Tris-HCl pH 7,5, 50% glycérol, 0,025 % (p/v) xylène cyanol et bleu de bromophénol) et déposé sur un gel natif 6% acrylamide/bisacrylamide (19/1, Rotiphorese Gel 40), et TBE 1x. Après migration à 150 V pendant 1h30, l'ADN est révélé au BET et la bande correspondant à l'insert est découpée, placée dans un boudin à dialyse contenant 300  $\mu$ L de TBE 1x stérile, et électroéluée pendant 30 min à 65 mA. L'éluat est récupéré, précipité à l'éthanol, lavé, séché et repris dans 10  $\mu$ L d'eau stérile.

#### **1.3.2. Digestion et déphosphorylation du vecteur pQE70**

Le vecteur pQE70 est digéré par les enzymes *Bam*HI et *Bgl*II. La digestion de 5  $\mu$ g de plasmide se fait pendant 2 h à 37°C dans un volume réactionnel de 100  $\mu$ L contenant le tampon *Bam*HI et 20 U de chaque enzyme. La réaction est arrêtée par une extraction phénolique suivie d'une précipitation à l'éthanol. La déphosphorylation est précédée d'un chauffage de 3 min à 90°C suivi d'un refroidissement brutal dans la glace pour éviter la recircularisation du vecteur qui présente des extrémités cohésives. L'ajout de 5 U de SAP pendant 30 min à 37°C permet de déphosphoryler les 2 extrémités 5' et 3'. La réaction est arrêtée par une extraction phénolique suivie d'une précipitation à l'éthanol. Le culot d'ADN est lavé, séché et repris dans 10  $\mu$ L d'eau stérile.

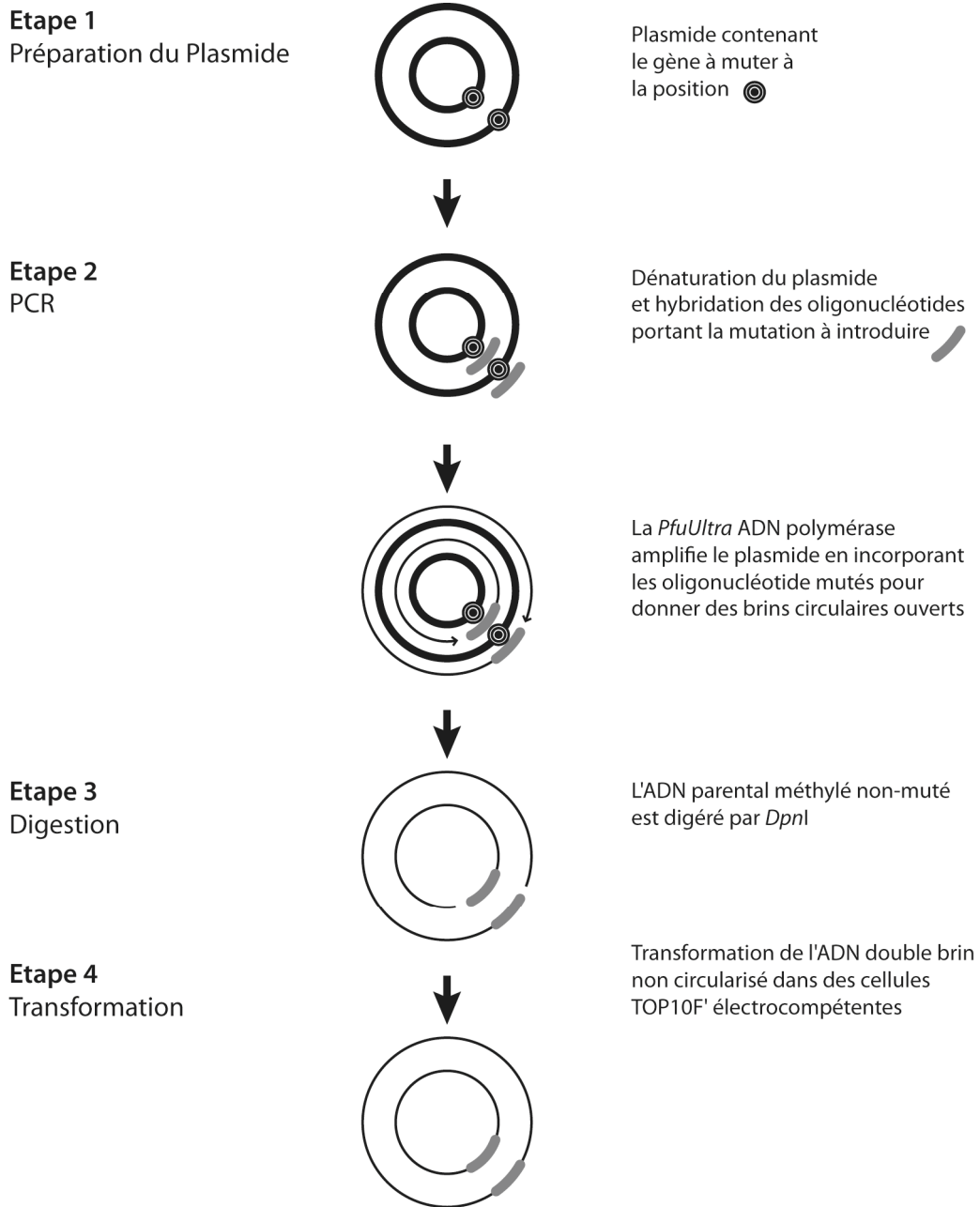
#### **1.3.3. Ligation**

La ligation de l'insert dans le vecteur digéré et déphosphorylé se fait dans les conditions décrites dans le paragraphe **C.3.** avec un chauffage préalable de 2 min à 90°C

suivi d'un refroidissement brutal dans la glace du vecteur pour faciliter l'insertion de l'insert. Une précipitation de l'ADN est nécessaire pour éliminer les sels avant l'électroporation. Le culot d'ADN est repris dans 10  $\mu\text{L}$  d'eau stérile.

#### **1.3.4. Transformation des cellules TG1 électrocompétentes**

La transformation de bactéries TG1 se fait par électroporation de 100  $\mu\text{L}$  de cellules en présence de 5  $\mu\text{L}$  de milieu de ligation à 2 500 V, 200  $\Omega$  et 25  $\mu\text{F}$ . Après ajout de 300  $\mu\text{L}$  de milieu LB aux cellules, 200  $\mu\text{L}$  sont étalés sur une boîte de Pétri contenant du milieu gélosé LB/Amp. La boîte est placée dans une étuve à 37°C pendant 16 h. Une dizaine de colonies sont prélevées et mises en culture dans 3 mL de milieu LB/Amp. La présence et l'orientation de l'insert sont vérifiées par digestion enzymatique.



**Figure 28 : Principe de la mutagenèse dirigée selon Stratagene**

Figure adaptée d'après le manuel QuickChange<sup>®</sup> Site-Directed Mutagenesis Kit de Stratagene.

## 2. Clonage des variants

### 2.1. La mt-TyrRS tronquée et le domaine C-terminal

Les gènes des deux variants de la mt-TyrRS : l'enzyme déletée de son domaine C-terminal (comprenant les résidus 1-347, mt-TyrRS- $\Delta$ S4) et le domaine C-terminal seul (résidus 348-453 mt-TyrRS-S4) sont clonés suivant la même procédure que le gène de l'enzyme complète en utilisant le vecteur pQE70 contenant le gène de la mt-TyrRS comme matrice. Les oligonucléotides utilisés pour la pCR sont les suivants : 5'-ATG CGA GGA TCC CAC TCG GGC GCT CAG GGG-3' et 5'-GGT GGT AGA TCT GTG ATA AAG GGC TTG TGT ACA CC-3' pour mt-TyrRS- $\Delta$ S4, 5'-GGA TCC GAT GCA CTG GAG GTC ATG TCC GAT-3' et 5'-AGA TCT CAA CTG AAG CCA TTT TAT AAT G-3' pour mt-TyrRS-S4.

### 2.2. Mutagenèse dirigée

L'introduction de mutations ponctuelles dans le gène de la mt-TyrRS ou de la mt-TyrRS- $\Delta$ S4 est possible par l'emploi de la technique de mutagenèse dirigée par PCR, résumée dans la **figure 28**. Pour chaque mutation le couple d'oligonucléotides utilisés doit répondre à plusieurs critères : (i) la région à muter doit être encadrée d'une quinzaine de nucléotides en 5' et en 3', (ii) les paires de bases distales doivent être des paires G-C et (iii) leur température de fusion doit être supérieure à 75°C. Dans un premier temps, une réaction de PCR est réalisée dans 50  $\mu$ L de milieu (20 mM Tris-HCL pH 8,8, 10 mM KCl, 10 mM  $(\text{NH}_4)_2\text{SO}_4$ , 2 mM  $\text{MgSO}_4$ , 0,1% Triton X-100 et 0,1 mg/mL de BSA) contenant 150 ng de chaque oligonucléotide, 0,2 mM de chaque dNTP, 10 ng de plasmide matrice et 2,5 U de Pfu Ultra® polymérase (Stratagene). Entre 12 et 18 cycles sont réalisés (30 sec à 95°C, 1 min à 55°C et 5 min à 72°C). Le brin parental méthylé est ensuite digéré spécifiquement par l'ajout de 10 U de l'enzyme de restriction *DpnI* pendant une heure à 37°C. L'ADN est ensuite précipité et repris dans 40  $\mu$ L d'eau. Enfin, 2  $\mu$ L sont utilisés pour transformer 50  $\mu$ L de bactéries Top10F' électrocompétentes.

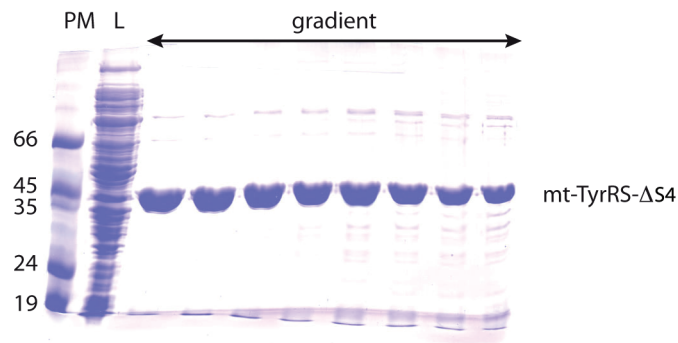
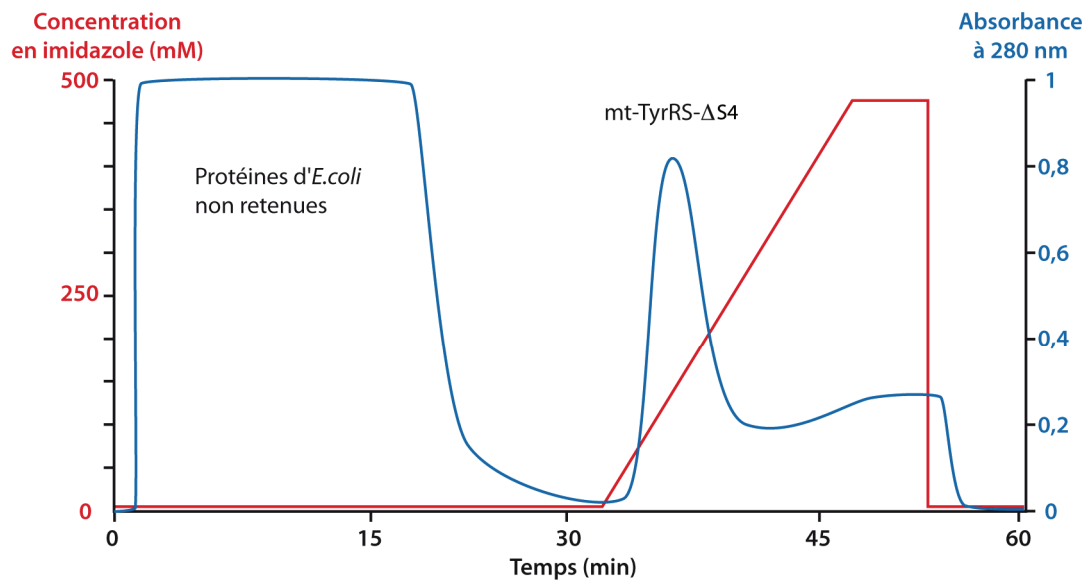
## 3. Expression

Le gène de la TyrRS mitochondriale et ses variants ont été clonés dans le vecteur pQE70 sous la dépendance du promoteur de l'opéron lactose. Une préculture de 5 mL de

bactéries Top10 transformées par le vecteur contenant le gène sert à inoculer 250 mL de milieu LB/Amp. Cette culture est mise à pousser à 37°C sous agitation jusqu'à atteindre la phase exponentielle de croissance ( $DO_{600}=0,7$ ). Les bactéries sont ensuite additionnées de 1 mM d'IPTG (qui stimule la transcription du gène). L'induction se fait pendant 16 h à 20°C sous agitation.

Pour la production de TyrRS- $\Delta$ S4 sélénométhionylée, l'incorporation de SeMet à la place de la Met dans la protéine recombinante est réalisée par culture d'une souche auxotrophe pour la Met dans un milieu où la Met est substituée par la SeMet. Cette substitution augmente le temps de génération des bactéries de 20 à 130 min. Pour cette raison, la culture est effectuée en deux étapes, la première visant à accumuler de la biomasse avant d'induire la production de la protéine sélénométhionylée. Ainsi, dans un premier temps, 500 mL de milieu M9 "modifié" additionné de 50 mg/mL de Met et des antibiotiques Amp et Cam sontensemencés avec 20 mL d'une préculture de bactéries RILX dans le même milieu. Cette culture est mise à pousser à 37°C sous agitation jusqu'à atteindre une  $DO_{600}=0,8$ . Les cellules sont alors centrifugées 5 min à 4 000 *g*, lavées deux fois avec 250 mL de milieu M9 (à 37°C) pour éliminer toute trace de Met et reprises dans 2 L du même milieu. Après 1h de culture à 37°C sous agitation, afin d'épuiser la Met endogène, le milieu est supplémenté de 36 mg/L de SeMet et les cellules sont remises à pousser jusqu'à atteindre la phase exponentielle de croissance ( $DO_{600}=0,5$ ). La production de la protéine recombinante est alors induite par l'ajout d'IPTG à la concentration finale de 1 mM pendant 16 h à 20°C.

Le gène de la TyrRS cytoplasmique humaine est cloné dans le vecteur pET20b sous le contrôle du promoteur de la T7 ARN polymérase. Dans ce cas particulier, l'addition d'IPTG permet d'induire la production de la T7 ARN polymérase (dont le gène est présent dans la souche BL21) qui à son tour va entraîner la transcription du gène de la TyrRS. La protéine est exprimée de la même façon que la mt-TyrRS.



**Figure 29. Purification par affinité de la mt-TyrRS- $\Delta$ S4 sur colonne NiNTA**

Le profil chromatographique est obtenu après culture de la souche bactérienne Top10 exprimant le vecteur pQE70-mtTyrRS- $\Delta$ S4. La concentration en imidazole est représentée en rouge et l'absorbance à 280 nm en bleu. Le gel de protéines correspondant comprend : le marqueur de poids moléculaire (PM, comprenant 5 bandes à 66, 45, 35, 24 et 19 kDa), un échantillon de lavage (L) et des fractions du pic d'absorbance.

## 4. Purification sur colonnes

### 4.1. Colonne d'affinité NiNTA

Les protéines surexprimées portent à leur extrémité C-terminale 6 résidus histidines. Cette queue poly-histidine a une affinité très forte pour l'ion Nickel ( $\text{Ni}^{2+}$ ), ce qui permet la rétention spécifique des protéines ainsi "marquées" sur une matrice couplée au Nickel. Grâce à cette stratégie, les différentes TyrRS ont pu être purifiées en une seule étape chromatographique. Toutes les colonnes utilisées sont installées sur un système de chromatographie à haute pression BioLogic Duo-Flow (Bio Rad).

Les cellules obtenues après induction sont centrifugées et le culot repris dans 100 mL de tampon A (50 mM  $\text{NaH}_2\text{PO}_4$  pH 8,0 et NaCl 300 mM) avec 5 mM de DTE par litre de culture, soniquées 6 fois 45 sec, puis centrifugées à 4°C pendant 30 min à 35 000 *g* pour éliminer les débris cellulaires. Le surnageant est récupéré et chargé sur une colonne de chromatographie contenant 2 mL de résine agarose-NiNTA (QIAGEN), équilibrée au préalable dans le tampon B (tampon A additionné de 20 mM d'imidazole). La majorité des protéines contaminantes sont éliminées au cours du lavage de la colonne avec 50 mL de tampon B. Les protéines retenues spécifiquement sur la matrice sont éluées en utilisant un gradient linéaire d'imidazole variant de 20 à 500 mM (tampons B à C). Tout au long de la chromatographie, le débit est maintenu à 2 mL/min. Les effluents sont collectés par fractions de 2 à 5 mL et l'absorption à 280 nm est enregistrée. Le profil d'absorption permet de suivre le déroulement des différentes étapes, et en particulier de détecter l'éluion de la protéine. Des aliquotes de 10  $\mu\text{L}$  des fractions d'éluion correspondant au pic d'éluion sont déposées sur un gel dénaturant de protéines pour en vérifier la pureté. Un exemple de purification de la mt-TyrRS- $\Delta\text{S4}$  d'après ce protocole est donné dans la **figure 29**.

La purification de la protéine en conditions oxydantes se fait en présence de DTNB à 5 mM (pour induire la formation de ponts disulfure (Riddles et al., 1983 ; Ivens et al., 2002) dans les tampons de sonication et de purification. Au cours de la purification sur colonne d'affinité, une étape supplémentaire de lavage avec 200 mL de tampon A contenant 5 mM de DTNB est réalisée avant l'éluion. Par la suite les protéines oxydées subissent les mêmes traitements que les protéines non traitées.

## 4.2. Colonne d'exclusion de taille

Ce type de matrice permet de séparer des macromolécules en fonction de leur encombrement. Pour analyser l'oligométrie de la mt-TyrRS, une colonne Bio-Prep SE-1000/17 est utilisée, équilibrée dans 50 mM HEPES-NaOH pH 6,5 et 300 mM NaCl. Un standard de protéines de poids moléculaires connus (Bio Rad) est chargé sur la colonne pour obtenir une courbe d'étalonnage. Des échantillons de 300  $\mu$ L sont chargés et le poids moléculaire apparent calculé d'après la courbe d'étalonnage.

## 5. Dialyses

Les fractions de purification contenant la cyt-TyrRS sont regroupées et dialysées contre un tampon de conservation (50 mM  $\text{KH}_2\text{PO}_4$  pH 7,2, 150 mM KCl, 10 mM  $\beta$ -ME, 50% glycérol) 16 h à 4°C, puis stockées à -20°C. Les fractions de mt-TyrRS, mt-TyrRS- $\Delta$ S4 et mt-TyrRS-S4 sont dialysées contre un tampon D (50 mM HEPES-NaOH pH 6,6, NaCl 300 mM, 20 mM  $\beta$ -ME) avec 50 % glycérol à 4°C pendant 3 h. Les échantillons utilisés pour les tests fonctionnels sont stockés à -80°C. Les échantillons destinés à la cristallogenèse sont dialysés une seconde fois contre le tampon D avec 10% de glycérol pendant 3h à 4°C. Enfin, la mt-TyrRS- $\Delta$ S4 est dialysée une dernière fois 16h contre le tampon D où les 20 mM de  $\beta$ -ME sont remplacés par 5 mM de DTE.



## E. Tests fonctionnels

### 1. Aminoacylation *in vitro* d'ARNt

#### 1.1. Principe du test d'aminacylation

La méthode d'aminacylation *in vitro* exploite le caractère acido-précipitable des ARNt, ce qui permet de séparer les acides aminés chargés sur un ARNt qui précipitent, des acides aminés libres solubles. L'utilisation d'acides aminés radioactifs permet de suivre et de quantifier la réaction d'aminacylation d'un ARNt.

#### 1.2. Les paramètres cinétiques étudiés

Selon les conditions expérimentales (concentration en ARNt, en aaRS et le temps d'incubation) trois paramètres de la réaction peuvent être déterminés : les constantes cinétiques  $k_{cat}$ ,  $K_M$  ainsi que les plateaux de charge.

Les plateaux de charge représentent l'équilibre entre la forme aminoacylée et non aminoacylée de l'ARNt. Cet équilibre dépend des vitesses (i) d'aminacylation, (ii) de déacylations spontanée et enzymatique et (iii) de la réaction inverse. Classiquement, les mesures de plateaux d'aminacylation sont réalisées avec 10-20 pmols d'ARNt et de 1 à 2  $\mu$ M d'enzyme pendant des temps relativement longs (10 à 30 minutes).

Les mesures cinétiques sont effectuées dans des conditions de vitesses initiales (le produit s'accumule de façon linéaire en fonction du temps). Les constantes  $K_M$  et  $k_{cat}$  sont déterminées à partir de la représentation en double inverse de Lineweaver et Burk :

$$1/v = K_M/V_{max} \times 1/[S_0] - 1/V_{max}$$

La constante de Michaelis-Menten ( $K_M$ ) reflète l'inverse de l'affinité, alors que la vitesse de la réaction est directement reliée au  $k_{cat}$  ( $V_{max}/[enzyme]$ ), et le rapport  $k_{cat}/K_M$  indique l'efficacité catalytique du système. Le rapport  $(k_{cat}/K_M)^{sauvage} / (k_{cat}/K_M)^{mutant}$  représente la perte d'efficacité catalytique (P) du mutant relative au système sauvage. Les conditions de vitesses initiales sont obtenues en présence de concentrations d'ARNt (sauvage ou mutant) de  $10^{-3}$  à 3 fois le  $K_M$  et de 100 à 800 nM de mt-TyrRS. Chaque expérience est reproduite au moins trois fois.

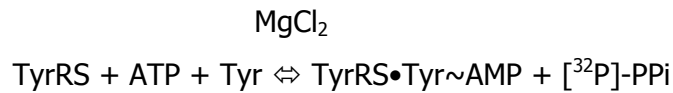
### 1.3. Le test d'aminocyclation

La tyrosylation est suivie *in vitro* en quantifiant la L-[<sup>3</sup>H]-tyrosine (49 Ci/mmol) ou la L-[<sup>14</sup>C]-tyrosine (465 mCi/mmol) fixée sur l'ARNt<sup>Tyr</sup>. Avant chaque réaction, les transcrits sont systématiquement renaturés dans l'eau stérile par un chauffage de 2 min à 65°C, suivi d'un refroidissement de 10 min à 25°C. L'enzyme est préparée dans un tampon de dilution (100 mM d'HEPES-NaOH pH 7,4, 1 mM de DTT, 5 mg/mL de BSA et 10 % de glycérol) de façon à obtenir un stock à 10 fois la concentration finale souhaitée. La réaction d'aminocyclation est initiée par l'ajout de 5 µL d'enzyme au milieu réactionnel (45 µL) préincubé à 37°C contenant 50 mM d'HEPES-NaOH pH 7,6, 25 mM de KCl, 12 mM de MgCl<sub>2</sub>, 2,5 mM d'ATP, 0,2 mg/mL de BSA, 1 mM de spermine, 10 mM de L-tyrosine radiomarquée et de l'ARNt<sup>Tyr</sup> à différentes concentrations. Les tests sont réalisés à 37° C dans un volume final de 50 µL, et des aliquotes de 15 µL sont prélevées à différents temps, puis déposées sur des rondelles de papier Whatman 3 MM. Ces rondelles sont plongées immédiatement et pendant 10 min dans une solution de TCA 5% froid dont le rôle est de précipiter l'ARNt. Cette solution de TCA est saturée en tyrosine froide, ce qui réduit la fixation aspécifique de tyrosine radioactive sur le papier Whatman. Les rondelles sont lavées deux fois au TCA 5%, puis rincées à l'éthanol et séchées. Après dépôt dans des fioles de comptage en verre contenant 2,5 mL de liquide de scintillation (Ecoscint O<sup>TM</sup>, National Diagnostics), la radioactivité fixée sur les rondelles est quantifiée dans un compteur à scintillation (LS6500, Beckman Coulter).

Lorsque la tyrosine tritiée est utilisée une correction des valeurs de comptage de l'activité spécifique (AS) de la tyrosine libre doit être apportée. Une partie des rayons β du tritium est arrêtée par le papier (contrairement au rayonnement <sup>14</sup>C). Cette extinction, appelée "quenching", est liée à la distribution différentielle de l'ARNt aminocyclé et de l'acide aminé libre dans le papier. En effet, l'ARNt précipité est retenu à la surface du papier, alors que l'acide aminé libre diffuse dans les fibres du papier. L'ampleur de cette extinction est mesurée en comparant des plateaux d'aminocyclation d'ARNt effectués exactement dans les mêmes conditions mais en présence de la tyrosine tritiée ou marquée au <sup>14</sup>C. Le nombre de pmoles aminocyclées étant identique, il est ainsi possible de quantifier le "quenching" de la tyrosine tritiée. Dans le cas de la tyrosine, le facteur de correction est de 2,8.

## 2. Echange ATP/[<sup>32</sup>P]-PPi

La réaction d'échange ATP/[<sup>32</sup>P]-PPi s'effectue à l'équilibre de la réaction d'activation de la tyrosine par la TyrRS :



Expérimentalement, on mesure la vitesse de formation d'ATP-[<sup>32</sup>P] à partir du [<sup>32</sup>P]-PPi introduit dans le milieu. La réaction s'effectue dans un volume de 200 µL contenant 100 mM d'HEPES pH 7,5, 10 mM de MgCl<sub>2</sub>, 2 mM de KF (pour inhiber la pyrophosphatase inorganique contaminante), 2 mM de [<sup>32</sup>P]PPi (environ 2 000 cpm/nmol), 2 mM d'ATP et 2 mM de tyrosine, à 37°C. La réaction est initiée par l'addition de 2 µg/mL de mt-TyrRS ou mt-TyrRS-ΔS4 préparée dans le tampon de dilution (100 mM d'HEPES-NaOH pH 7,4, 1 mM de DTT, 5 mg/mL de BSA et 10 % de glycérol). Des aliquotes de 50 µL sont prélevées à différent temps de la réaction, mélangées à 200 µL d'une suspension de charbon activé 1% (p/v) dans 0,4 M de pyrophosphate de sodium et 15% (v/v) d'acide perchlorique, inactivant l'enzyme et arrêtant la réaction. L'ATP-[<sup>32</sup>P] formé s'adsorbe, par le cycle de l'adénine, sur le charbon activé. La suspension est filtrée sur un disque en fibre de verre (GFC, ø 2,5 cm, Whatman), placé dans un bloc branché à une trompe à vide et immobilisé par un cylindre métallique percé d'un puits au diamètre du filtre. Le charbon retenu sur le filtre est lavé trois fois avec 5 mL d'eau déminéralisée, puis rincé avec 5 mL d'éthanol et séché. La radioactivité de l'ATP-[<sup>32</sup>P] retenue sur le disque est mesurée dans des fioles en verre en présence de 2,5 mL de liquide scintillant.

## F. Cartographie des ARNt en solution

### 1. Principe

La caractérisation de la structure de l'ARN en solution est basée sur l'utilisation de sondes structurales chimiques ou enzymatiques reconnaissant une conformation de l'ARN (simple-brin ou appariée) avec une spécificité de nucléotide ou non (Brunel et Romby, 2000; Giegé et al., 2001). Les réactions sont optimisées pour produire au plus une coupure (ou une modification) par molécule d'ARN. La réaction est révélée à l'aide d'oligonucléotides radio-marqués séparés sur gel de polyacrylamide.

### 2. Marquage des ARNt en 5'

Les acides nucléiques sont marqués à leur extrémité 5' dépourvue de groupement phosphate grâce à la T4 polynucléotide kinase (T4 PNK). Une étape de déphosphorylation préalable des ARNt issus du transzyme par la SAP est donc nécessaire. La déphosphorylation de 5 µg d'ARNt est réalisée à l'aide de 2,5 U de SAP dans un tampon SAP (10 mM Tris-HCl pH 7,5, 10 mM MgCl<sub>2</sub> et 0,1 mg/mL BSA) pendant 30 min à 37°C. La réaction est arrêtée par une extraction phénolique. Les acides nucléiques sont précipités et repris dans un volume de 10 µL.

Le marquage en 5' de 80 pmoles d'ARNt est effectué pendant 30 min à 37°C en présence de 20 µCi de [<sup>32</sup>P]-ATP, dans 15 µL d'un milieu comprenant 50 mM Tris-HCl pH 7,6, 10 mM MgCl<sub>2</sub>, 5 mM DTE, 1 mM ATP et 10 U de T4 PNK. Après ajout d'un volume de tampon de dépôt, le milieu est soumis à une électrophorèse sur gel analytique 8% acrylamide/bisacrylamide (19/1, Rotiphorese Gel 40), urée 8 M pendant 1 h à 150 V.

La bande correspondant à l'ARNt radiomarqué est révélée par autoradiographie, et découpée. L'ARNt est récupéré par élution passive dans 400 µL de tampon d'élution 1 h 30 à 25°C puis précipité par l'ajout de glycoène et repris dans un volume d'eau permettant d'obtenir 70 000 cpm/µL (mesure Cerenkov).

### 3. Cartographie

La structure de l'ARN peut être caractérisé au moyen de sondes enzymatiques, des endoribonucléases (RNases) qui reconnaissent spécifiquement les régions simple-brins (RNases S1, T1, T2) ou les régions appariées et structurées (RNase V1).

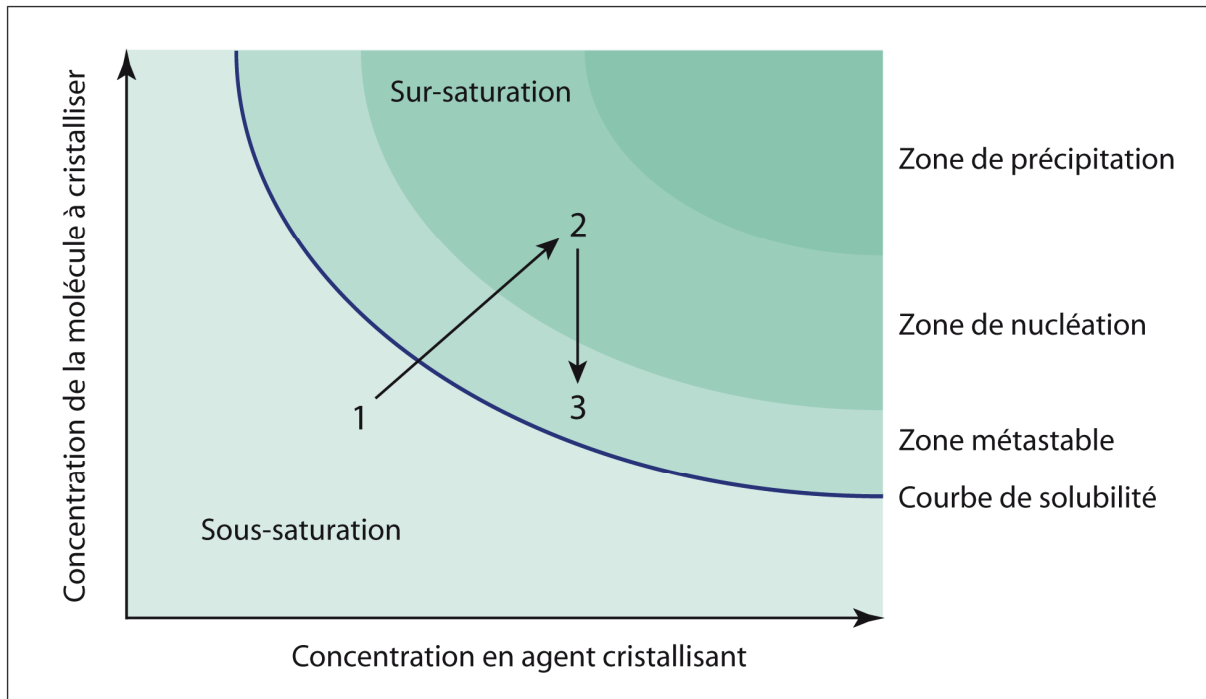
Les ARNt sont dans un premier temps dénaturés/renaturés par chauffage à 60°C pendant 2 min dans l'eau suivi d'un refroidissement lent. Le milieu réactionnel de 20 µL contenant le tampon (50 mM HEPES-NaOH pH 7,5, 10 mM MgCl<sub>2</sub> et 25 mM KCl), 70 000 cpm d'ARNt marqué, 1 µM d'ARNt non marqué est incubé 5 min à 25°C. La sonde enzymatique est ajoutée (16 mU de nucléase T1, 12 mU de nucléase T2, 0,18 U de nucléase V1 ou 1 U de nucléase S1 et 1 mM de ZnCl<sub>2</sub>) et le milieu est incubé 10 min à 25°C. La réaction est arrêtée par addition d'un volume de stop mix (0,6 M AcONa pH 5, 3 mM EDTA, 0,1 mg/mL d'ARNt total d'*E. coli*). Les produits de clivage sont purifiés par extraction phénolique puis précipités à l'éthanol. Afin de visualiser d'éventuelles dégradations de l'ARNt au cours des manipulations, des expériences de contrôle sont également effectuées sans sondes enzymatiques.

L'échelle alcaline est réalisée dans 10 µL de milieu réactionnel contenant 35 mM de NaHCO<sub>3</sub> pH 9, 70 000 cpm d'ARNt marqué, 1 µM d'ARNt non marqué pendant 4 min à 80°C puis plongé dans la glace. L'échelle de guanosine est effectuée dans 10 µL de milieu contenant 10 mM de citrate de sodium pH 4,5, 3,2 M d'urée, 0,4 mM d'EDTA, 70 000 cpm d'ARNt marqué, 1 µM d'ARNt non marqué. Le milieu est incubé 5 min à 60°C puis additionné de 2 U de ribonucléase T1 et incubé 4 min à 60°C. La réaction est arrêtée dans la glace. Les produits de clivages enzymatiques et les 2 échelles sont additionnés de 0,2 volumes de glycogène (1/1) et précipités à l'éthanol avant d'être repris dans 10 µL de solution de dépôt.

La séparation et la visualisation des produits de clivage se fait sur gel de polyacrylamide en conditions dénaturantes. La migration des fragments en présence d'urée et à haute température permet un dépliage des molécules d'ARN et donc une séparation selon leur taille sur gel de polyacrylamide. La migration est réalisée sur un gel de polyacrylamide 12% (19/1, Rotiphorese Gel 40) avec de l'urée 8 M à 1 600 V, 55 mA et 50 W jusqu'à ce que le bleu de bromophénol soit à 10 cm du bas de la plaque. Le gel est ensuite démoulé, transféré sur une autoradiographie usagée (soutient le gel) et recouvert d'un film "saran wrap". Le gel est conservé ainsi à -80°C et révélé par autoradiographie.

# G. Cristallographie

## 1. Principe de la cristallogenèse



**Figure 30. Diagramme de solubilité théorique**

Le diagramme décrit le comportement d'une molécule en fonction de sa concentration et de la concentration en agent cristallisant. D'après Riès-Kautt et Ducruix, 1992.

La cristallographie est une technique de choix pour l'étude de la structure tridimensionnelle de macromolécules biologiques (Rossmann et Arnold, 2001). Un cristal est un empilement périodique d'une molécule dans les trois dimensions de l'espace. Néanmoins, l'obtention d'un cristal ordonné reste un facteur limitant de cette méthode. La cristallisation a donc pour but de déplacer une molécule en solution vers un état de sursaturation propice à la nucléation (passage de 1 à 2 dans la **figure 29**). À partir du noyau, le cristal va croître de manière régulière et ordonnée jusqu'à rejoindre un état d'équilibre par épuisement des molécules en solution (passage de 2 à 3). Les conditions de cristallisation impliquent de nombreux paramètres (concentration, pH, température, agent cristallisant, additif ...) qui diffèrent pour chaque molécule étudiée. Il est donc nécessaire de tester un grand nombre de combinaisons de ces paramètres pour obtenir des cristaux.

## 2. Cristallisation par diffusion de vapeur

Différentes techniques de cristallisation ont été mises au point et sont couramment utilisées (par exemple diffusion de vapeur, batch, dialyse, revues dans (Ducruix et Giegé, 1999)). Parmi elles nous avons utilisé la diffusion de vapeur qui consiste à mélanger dans une goutte la solution de protéine et d'agent cristallisant avant de la placer dans une enceinte close contenant une solution d'agent cristallisant concentré en excès. La concentration de l'agent cristallisant et de la protéine dans la goutte va alors augmenter par évaporation jusqu'à arriver à l'équilibre avec la solution concentrée en agent cristallisant, ce qui permet d'amener progressivement la protéine en conditions de sursaturation.

Pour la recherche des conditions de cristallisation de la TyrRS mitochondriale sous ses différentes formes, plusieurs kits commerciaux ou mis au point au laboratoire ont été utilisés (Classics, PEGS, SM1, SM4, AmSO4 de Nextal et Index de Hampton Research), de manière à tester un grand nombre de combinaisons de tampons, d'agents cristallisants et d'additifs. De plus l'utilisation d'un robot de cristallisation nanogouttes (Mosquito® , TTP Labtech Inc., UK) permet de réduire le volume de solution de protéine utilisé pour chaque goutte jusqu'à 100 nL, ce qui permet de multiplier les essais en jouant notamment sur la concentration en protéine dans la goutte. Le format des boîtes de cristallisation utilisées est celui de la marque Greiner 96 puits qui permet de réaliser 3 gouttes par condition par la méthode de diffusion de vapeur en gouttes assises. La cristallisation a aussi été testée en présence de petits ligands utilisés individuellement ou de manière combinée à une concentration finale de 5 mM : tyrosine, ATP, AMPpCp, YSA et TYA. Les essais de cristallisation du complexe ARNt/synthétase ont été faits en utilisant l'ARNt<sup>Tyr</sup> natif d'*E. coli* dans un rapport 1,5 : 1 et après vérification de la présence d'un complexe homogène par DLS. L'optimisation des conditions de cristallisation a été faite avec le kit Opti-Salts de Nextal qui permet de tester une gamme de sels et de pH. Tous les cristaux ont été obtenus à 20°C.

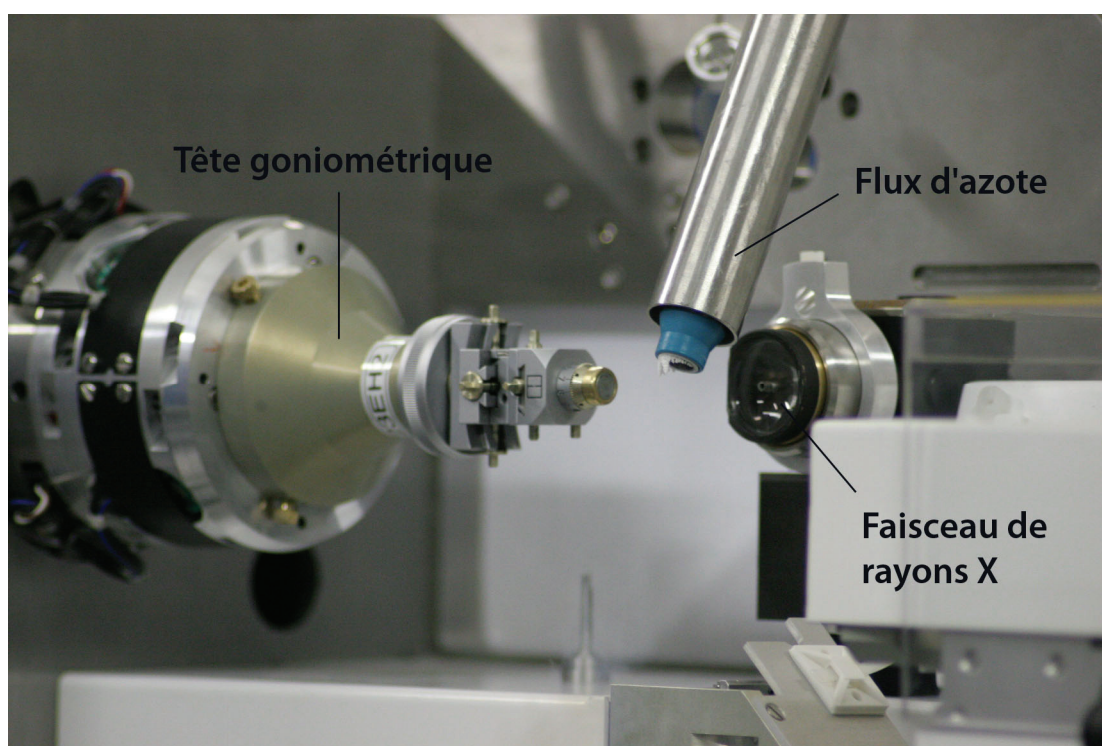
## 3. Diffraction des rayons X

L'agencement ordonné des macromolécules au sein d'un cristal lui confère des propriétés de diffractions des rayons X qui permettent de déduire la structure de la macromolécule élémentaire d'après les clichés de diffractions générés (Rossmann et Arnold, 2001). Les différents cristaux de TyrRS mitochondriale obtenus ont été analysés sous

rayonnement X synchrotron à l'ESRF (European Synchrotron Radiation Facility) de Grenoble. Un exemple de dispositif expérimental utilisé est présenté dans la **figure 30**. Les cristaux sont montés sur des boucles en nylons fixées à une tige métallique de 18 mm puis plongés dans une solution cryoprotectante de paratone avant d'être montés sur la tête gonométrique et congelés sous le flux d'azote qui protège le cristal des dommages engendrés par le rayonnement X. Le capteur CCD disposé derrière le cristal permet d'enregistrer et de mesurer l'intensité des taches générées par la diffraction du rayon à travers le cristal. La distance entre le cristal et le détecteur est ajustée en fonction de la limite de diffraction du cristal étudié. Pour collecter un jeu de données complet le cristal tourne dans le faisceau incident par pas d'un demi à un degré autour d'un axe perpendiculaire au faisceau incident. À chaque oscillation, un cliché de diffraction est enregistré. L'oscillation totale nécessaire pour collecter un jeu complet est définie en fonction des symétries existant dans le cristal.

Quand on ne dispose pas de protéine homologue ou de structure proche de celle de la protéine cristallisée on utilise la méthode de dispersion anormale. La dispersion anormale est la propriété de ré-émission des photons avec la même énergie que la diffusion normale. Le sélénium, Se, est l'un des atomes les plus utilisés en dispersion anormale car il peut être substitué aux atomes de soufre des protéines. Les positions des diffuseurs anormaux sont déduites en travaillant à plusieurs longueurs d'ondes. Une fois les positions connues, les phases peuvent être calculées. Pratiquement, un spectre de fluorescence des cristaux de protéine sélénométhionylée est d'abord enregistré afin de vérifier la présence de signal anormal dû au sélénium. Puis le cristal est soumis à trois collectes de jeux de données qui diffèrent par la longueur d'onde du faisceau correspondant à celles du pic, du point d'inflexion et du remote du spectre.





**Figure 31. Dispositif expérimental utilisé pour l'enregistrement de clichés de diffraction**  
(A) L'appareillage est installé dans une pièce où la température et l'humidité sont contrôlées. (B) Le cristal est monté sur la tête goniométrique de façon à être dans l'axe du faisceau de rayonnement X et maintenu congelé par le flux d'azote. Après montage du cristal, l'ensemble de l'appareillage est piloté à distance. Photos prises le 16 Juin 2006 sur la ligne microfocus ID23-2 de l'ESRF, Grenoble.

## 4. Résolution de structure

Les jeux de données enregistrés sont traités informatiquement afin de déterminer la géométrie de la maille élémentaire du cristal, d'indexer les tâches de diffraction et d'en déterminer les intensités. La suite de programmes *HKL* et sa version interfacée *HKL2000* (Otwinowski et Minor, 1997) a été utilisée à cet effet. L'orientation et les paramètres de maille du cristal sont déterminés à partir du premier cliché de diffraction avec le programme *denzo*. Les tâches sont ensuite indexées de manière récurrente et leur intensité mesurée sur l'ensemble du jeu d'images. Les réflexions sont ensuite mises à l'échelle par le programme *scalepack*. Pour pouvoir calculer une carte de densité électronique, il est nécessaire de phaser les données. Pour la TyrRS mitochondriale, le remplacement moléculaire (MR) et la diffusion anormale à plusieurs longueurs d'onde (MAD) ont été utilisées. Le remplacement moléculaire consiste à placer dans la maille cristalline la structure d'une protéine homologue comme modèle de manière à obtenir un premier jeu de phases. Dans le cas présent, il s'agit de la structure de la TyrRS bactérienne d'*E. coli* (référence pdb 1VBM ; Kobayashi et al., 2005). La recherche de solutions a été réalisée avec le programme *Molrep* de la suite *CCP4* (Collaborative Computational Project, 1994). Une carte de densité est alors calculée, qui va s'améliorer au fur et à mesure de la reconstruction de la TyrRS mitochondriale. L'affinement du modèle alterne construction dans la carte de densité et affinement par minimisation d'énergie ou dynamique moléculaire. Son évolution est suivie par le biais du facteur  $R$  d'accord cristallographique et du facteur  $R_{libre}$  (calculé avec un jeu test comprenant 5% des données de départ et exclues des calculs) qui convergent de façon concertée. Les modèles ont été construits avec le programme *Coot* (Emsley et Cowtan, 2004) et l'affinement effectué avec les programmes *refmac* (Collaborative Computational Project, 1994), *cns* (Brünger et al., 1998) et *phenix* (Adams et al., 2002) Enfin la qualité du modèle est évaluée à l'aide du logiciel *procheck* (Laskowsky et al., 1993).

## H. Bioinformatique

### 1. Alignements de séquences

Les alignements de séquences ont été principalement réalisés à distance à l'aide des programmes mis à disposition par la communauté scientifique sur nombreux servers, notamment l'EBI (European Biology Institute, UK) pour le programme *ClustalX* (Thompson et al., 1997) et l'IGS (Institut de Génomique Structurale, Marseille) pour les programmes *Tcoffee* (Notredame et al., 2000) et *3Dcoffee* (O'Sullivan et al., 2004). Les ajustements manuels des alignements ont été faits avec les programmes *ClustalX 1.82*, *Seaview* (Galtier et al., 1996), *BioEdit*, *MEGA4* (Tamura et al., 2007) ou des éditeurs de texte (e.g. *SubEthaEdit* sur Mac OS X ou *Crimson Editor* sous Windows XP).

### 2. Arbres phylogénétiques

L'alignement des séquences préalablement réalisé avec *Tcoffee* (Notredame et al., 2000) a été édité manuellement avec le logiciel *BioEdit* pour restreindre l'alignement au domaine catalytique et supprimer les gaps. Le modèle de protéines a été sélectionné avec le logiciel *MODELGENERATOR*, version 0.82 (Keane et al., 2006). Le modèle proposé est le WAG (Whelan et Goldman, 2001). L'arbre Neighbor-Joining (NJ) a été construit en utilisant le logiciel *Phylo\_win*, version 2.0 (Galtier et al., 1996) avec des distances de Poisson. L'analyse par Maximum Likelihood (ML) a été faite avec le logiciel *PHYML*, version 2.4.4 (Guindon et Gascuel, 2003) en utilisant le modèle d'évolution de protéines WAG. Les valeurs de bootstrap sont de 500 répliquats pour l'analyse NJ et 100 répliquats pour l'analyse ML.

### 3. Outils bioinformatiques

De nombreux outils informatiques ont été utilisés au cours de ce travail, notamment *Esript* (Gouet et al., 1999) pour la mise en page des alignements de séquence, *Endscript* (Gouet et Courcelle, 2002 ; Gouet *et al.*, 2003) pour la recherche des contacts cristallographiques, *Pymol* (DeLano, 2002) pour la réalisation des figures avec le module

*APBS Tools* (Baker et al., 2001) pour le calcul des potentiels électrostatiques. La recherche de nouvelles séquences de TyrRS mitochondriales a été principalement faite à partir de la base de données *Ensembl*. La prise en main des différents outils a été facilitée par la lecture de l'ouvrage *Bioinformatics for Dummies* (Claverie et Notredame, 2006). Enfin, l'ensemble des figures a été réalisé avec la suite de logiciels du pack Adobe® Creative Suite 2 (Photoshop, Illustrator, In Design).

# Annexes



## Articles

- 1- L. Bonnefond, M. Frugier, M. Sissler, C. Florentz, R. Giegé et J. Rudinger-Thirion (2004) *Evolution des systèmes d'aminacylation spécifiques de la tyrosine*, Congrès International de Biochimie, Marrakech, Maroc, **Actes du Congrès**, Enzymologie et Métabolisme (M. Baaziz & A. Hakkou, eds.), pp. 109-112
- 2- L. Bonnefond, A. Fender, J. Rudinger-Thirion, R. Giegé, C. Florentz et M. Sissler (2005) *Toward the full set of human mitochondrial aminoacyl-tRNA synthetases: characterisation of AspRS and TyrRS*, **Biochemistry** 44, 4805-4816
- 3- L. Bonnefond, M. Frugier, R. Giegé et J. Rudinger-Thirion (2005) *Human mitochondrial TyrRS disobeys the tyrosine identity rules*, **RNA** 11, 558-562
- 4- L. Bonnefond, R. Giegé et J. Rudinger-Thirion (2005) *Evolution of the tRNA<sup>Tyr</sup>/TyrRS aminoacylation systems*, **Biochimie** 87, 873-883
- 5- L. Bonnefond, M. Frugier, E. Touzé, B. Lorber, C. Florentz, R. Giegé, J. Rudinger-Thirion et C. Sauter (2007) *Tyrosyl-tRNA synthetase: first crystallization of a human mitochondrial aminoacyl-tRNA synthetase*, **Acta Cryst.** F63, 338-341
- 6- L. Bonnefond, M. Frugier, E. Touzé, B. Lorber, C. Florentz, R. Giegé, C. Sauter et J. Rudinger-Thirion (2007) *Crystal structure of human mitochondrial Tyrosyl-tRNA synthetase reveals idiosyncratic features*, **Structure** (soumis)

## Présentations orales

### *Évolution des systèmes d'aminacylation spécifiques de la tyrosine*

- Forum des Jeunes Chercheurs, **Marrakech** (Maroc), 2-6 Mai **2004**
- 5<sup>ème</sup> rencontre sifrARN, **Arcachon**, 10-13 Octobre **2004**
- Séminaire interne IBMC, **Strasbourg**, 16 Septembre **2005**

### *Evolution of tyrosine specific aminoacylation systems*

- Tokyo Institute of Technology, **Tokyo** (Japon), 26 Septembre **2005**
- RIKEN Yokohama, **Tokyo** (Japon), 28 Septembre **2005**
- RIKEN Wako, **Tokyo** (Japon), 30 Septembre **2005**
- Seoul National University, **Seoul** (Corée), 21 Octobre **2005**

### *Structural and functional properties of human mt-TyrRS deprived of its C-terminal domain*

- International Conference on Aminoacyl-tRNA Synthetases 2006  
**San Diego** (CA USA), 1-6 Octobre **2006**

### *Caractérisation structurale d'une forme minimale de la TyrRS mitochondriale humaine*

- Journée Scientifique de l'UPR 9002, **Strasbourg**, 14 Décembre **2006**

## Affiches

L. Bonnefond, M. Frugier, M. Sissler, R. Giegé, et J. Rudinger-Thirion (**2003**) *Human mitochondrial tyrosyl-tRNA synthetase: expression, purification and activity*

International tRNA workshop, **Banz** (Allemagne), 2-7 Octobre 2003

L. Bonnefond, R. Giegé et J. Rudinger-Thirion (**2005**) *Le système tyrosine fait entorse aux règles d'aminacylation*

Congrès annuel de la SFBBM et Forum des Jeunes Chercheurs, **Nantes**, 24-26 Octobre 2005

L. Bonnefond, M. Frugier, R. Giegé, et J. Rudinger-Thirion (**2005**) *Mitochondrial TyrRS charges eukaryotic tRNA<sup>Tyr</sup>*

21<sup>st</sup> International tRNA Workshop, **Bangalore** (Inde), 2-7 Décembre 2005



# Séquences des TyrRS mitochondriales

Le format utilisé est le suivant :

>nom\_de\_l'organisme-mito

Séquence

Taille en acide aminés / premier acide aminé après le signal d'adressage / probabilité d'adressage à la mitochondrie prédite avec l'algorithme MITOPROT (Claros et Vincens, 1996) / commentaires

>Aedes.aegyptis-mito

MILRRVVICLVRPAVVQLRRTYTEGNILKLRDRGFFQDVFPAEAIDKARSMLGASPQTVY  
AGFDPTADSLHVGNNLLVLIIGLLHAQRAGHQPIALLGGATGQIGDPSGRKTERQAMLLDMV  
NHNLACIGSQMDTIFQNRKFLWEKRGHSGALKPVLVNNADWYRELNFVQFISVGRHF  
RMGSMLSRASVQSRLQSEGGMSFTEFTYQIFQAYDWLHLLKAHNCRFQLGGSDQMGNIMS  
GHELISRVDKDVYGLTLPITNEEGDKFGKSAGNAVWLSGDKTSPYALYQFFVVRTPDSE  
VEKLLKLLTFMPLDEIRSVMARHQRTPELWEAQKLLAQELTLLVHGEEGLKKAEDITAAL  
YSGNVEALGELEVKDIKQTFGGAPVVDILPEPGMTVLQVAMRAKCFPTEKDAIRIIGAGG  
FTINLNKAKNPAEVLAPGVHILKNGISLLRVGKRNYIIVKWL

463 aa / 14e aa / 0,9304

>Anopheles.gambiae-mito

MLASTRTFQLKARLKTLLRQPARWYTERNILKLRDRGFIQDVFPAETIDKASSMLGSTCQ  
TVYAGFDPTANSLHVGNNLLVLIIGLLHTQRAGHQPIALVGGATGLIGDPSGRKTERQQLET  
EAVEHNVT CIRRIETIFANHERFFWDRPSSLKPVLLVNNADWYRQYSFVDFMANVGRHF  
RMGAMLSRSSVQTRLTSESGMSFTEFSYQLFQAYDWLHLLRQHNCRFQLGGSDQMGNIMS  
GQELISRTDGKEVFGLTLPVLTNEEGDKFGKSAGNAVWLSDDRTSPYAMYQFFVVRTPDSE  
VEKLLRLLTFLPVQTI EQAMARHRRTPELWEAQKLLAGELTKLVHGEAGLYKAMAI SRAL  
YNGDLSTLEQLEVRDVAQSFQGGAPLCEILPEPGMTVLDVALRARCFFPSRADAERIIGAGG  
FSINLKKAKNPAEVLSPSVHILSNRISLLRVGKRNYIIVKWL

462 aa / 21eme aa / 0,9893

>Apis.mellifera-mito

MNNLMRLRPFACKLNSRKLHFAKYILDENKQLYEDVFPDICMNDIKKLWKSPPQCIYAG  
FDPTAESLHIGNLLILINLLHWQRCGHEVI ILLGGATGLIGDPSYRKKERVEIDQIILKE  
NLECI RNDIKTIIRNHTQYFCKNHNHLIPIKILNNLQWYKDMNLLSFIKDVGKYFRMGTM  
LGRTSVRARLES DNMSFTEFTYPIFQGYDWLHLQRTYCNCFQVGGQDQMGNISGYDLV  
TKYTKNQVYGLTLPITAEAGGKFGKSTGNAIWLSPAKSSSFQLYQYFMRVEDADVEKFL  
HFFTFPLPVKIKKIIIEHFKKPELRI AQRNLAEKVTILVHGEAGLIAAKRASAILYDKSI  
DSLAKM NANELVQILEGANLIELLSEPGINVEYELAMKANCFFKSDYDARRIIEAGGFYINY  
QKITNMEEVVVPGIHILN NNISLLRVGKKTTHIVHWI

457 aa / 18eme aa / 0,9773

>Arabidopsis.thaliana-mito

MAYATGITFASRSILPICSRFTLSPLRVASLLVFPEKSSATFFRRVQVPHLFSTSTTTLF  
SSVKCSIHSTSSPETENQAVFRPNVVDILEERGLLESITSENLRSAACSDPKVAPLRVYCG  
FDPTAESLHLGNLLGIIVLSWFQRCGHQAVGLIGGATGRVGDPSGKSLERPELDADTLEK  
NIAGITKIIKILGSNPSPGGSYVIFNNYDWWKDMTMLDFLNKVGRFARVGTMMAKESVK  
KRLESEQGMSTYFTYQLLQAYDFLHLFKNEGIVQIGGSDQWGNITAGTDLIRKILQAE  
EAAYGLTFPLLLKNDGTFKFKSEGDGAIWLSPSMSPYKQYQYFFSVPDQVIRFLKTLTF  
LSLDEIKILEDQMSKPGYVNPNTAQIKLAAEVTRFVHGEEGLKEAIKATEALRPGAETKLD  
WNLIERIAEDIPSCSLPIDRVSGLSIVDLSVSAGLFESKSAARRMLKQGGFYMNNERVDD  
ENKRVEEDIVEGRGLVLSAGKKNKVVVRIS

511 aa / 47eme aa / 0,9769

>Bos.taurus-mito

MAAPMLRCISRSQWLGTQALSGVLPLGLREAHSAAQGLLAVQKARGLFKEFFPEKGTKTE  
LPELFDRGTGGKFPQTIYCGFDPTADSLHVGHLLALLGLFHFQORAGHNVIALVGGATARL  
GDPSGRTKEREALDAERVQSNARALSQGLKALAAHQQLFANGRTWGSFTVLDNSAWYQK  
QDLVNFLLAAVGGHFRMGTLTLLSRLSVQTRLKSSEGMSLAFLYQVQLQAYDFYYLFQHYGCR  
VQLGGSDQLGNIMSGYEFIHKLGTGEDVFGISVPLITSTTGAKLGKSAGNAVWLNDRDKTSP  
FELYQFFLRQDDLVERYLKLFTFLPLPEIDHIMQLHVKEPEKRGPKRLAAEVTKLVHG  
QEGLASAKRCTQALYHSSIDALEVMSDQELKELFKEASFSELVLDPGTSLVLDTCRKANAI  
PDGPRGYRMITEGGVSINHRQVTNPESVLVVGQHILKNGLSLLKIGKRNFYIIKWLRL

478 aa / 17eme aa / 0,9276

>Candida.albicans-mito

MLKTHTRNIPLIRRLARFNSTIARDPVTIIPTIYELTEAKDLTPETNPDNSLLEYLQSRH  
LIESITDDNLYKLTKRGSNHKFKLYCGADPTASSLHLGNLLPLMVLLHFKMSGNDVVGLV  
GGATGLVGDPSGRKTERNKIDEEVEEDNVTKIQRQISTFLSNGIEYAKSRQFPMTEKVG  
TTSVNNASWVWENKMLEFLATYGRHIRVSSMLARDSIQSRLELGGIGFNEFTYQILQAYD  
FWHLYKDENVNMVQVGGNDQWGNITAGIDLISRLKKFHGKKHEAYGLTVPLLTSSGKFG  
KSAGNAVFIDSSLTTPYQMYQYFINVPDDIVGKLLKVFTLLPLNVI EGELLPKHNSDPGL  
RIAQRVLAREVVDLIHGVGVGDEMAFITGFLFPTPDQPFNDNVSADKLIENFKRSGILFK  
RNKPSPEEIKLSALLADIVGKSKGEMRRLIKAGGVYMGDRNQIEDPDDVILFDVDNHL  
IDGKLLLVRVGKQNYVVEFV

501 aa / 26eme aa / 0,9458

>Caenorhabdiis.briggsae-mito

MFLSKWMDAAIKTVRSPRANPFLDYIKDLDRKPLKHSYPADLLTKFDYELKQIQPYLY  
AGFDPTAESLHVGNLLILLNLIRSQQFGLRPIALIGFTASIGDPSGKKSERDLLPSDVI  
VSNAQSVTEQVRRIFANASSDTEPPIIVNNNDWLEKAALRDFFRECKSMQLGKMLRMKTI  
KNRLEKGLSYTEFSYQTMQAYDWYTLSEKFGCRFQLGGYDQLGHLDFGAHYIKKRCNQSQ  
GTDNGFAAGVCFPIILTDSAGNKGKSEGGALWLDPKKTS PFHLYQFFAQLHDDKAEELL  
LLLSLRNFEDLESSLVKHKNLGKWI VQTELAMEITKVVHGEEGLDTALRCTKAMFGSKK  
ADLTGLSKEEVLQMFRTTVDLKKEEVISMGKLADLTRFGSGKGHLLMKNGAFSVNGAKKV  
DPTESINLLSVLPELEDLTLIRWGKRQYQVLRWI

455 aa / 21eme aa / 0,6695

>*Ceanorhabditis.elegans*-mito

MFNSKRIGNVALKTVRAPRESSFVDYITDLNARKQLQHSYPTDLLSKCSEDLRQLPPYVY  
AGFDPTAESLHIGNLLILVNLIRAQQFGIRPIALIGEFTASIGDPSGKKSERGLLAGDVI  
MHNSRKVTDQICKIFENAPGSSEKPIIVNNNDWLGKISLRDFLRECKNMQVGKMLRMNTI  
KNRLEVGLSYTEFSYQTMQAFDWTYTLSEKYGCRFQLGGYDQLGHLDFGAHYIKKMMNQAF  
AAGVCFPILTDSTGAKLGKSEGGALWLDATKTSPPHFYQFFAQLHDDKAEELLLLFSLQ  
DIEHIRDVLKNHRSNLGQWIAQRELA AEITRIVHGKEGLEVAMRCKAMFGAKKADLSGL  
SRSEVLQLFRTTIDLKKNENVATMGQLADASRLGSGKGHLLMQQGAFSVNGEKKRSPSESI  
ADVFLNASDLTLVCWGKRGYQLVRWV  
447 aa / 21eme aa / 0,8769

>*Canis.familiaris*-mito

MAAPMLRRFPRGRRFGTGPVSGVLLLLGQREARSGAQGLLAVQKARGLFKEFFPEKGTKVE  
LPELLDRSGSFPQTVYCGFDPTADSLHVGHLTLLGLFHFQ RAGHNVIALVGGATARLG  
DPSGRPKEREALAPEQVRGNARALRRGLEAVAANHQRLFANGRPWGSFTVLDNAAWYQGO  
HLVDFLAAVGGHFRMGTTLSRASVQARLQSAGGMSLAEFLYQVLQAYDFYYLFOHYGCRV  
QLGGSDQLGNITSGYEFIHKLTGEDVFGITVPLVTSTTGAKLGKSAGNAVWLNREKTSPP  
ELYQFFVRQQDNLVERYLKLFTFLPLPEIDHIMQLHDKEPEKRGPKRLAAEVTKL VHGO  
EGLDSAKRCTQALYHSSIDALEVMSDQELKELFKEASFSELVLDPGTSLVDICRKANAI P  
DGPRGYRMITEGGVSVINHRQVTNPESVLVVGQHILKNGLSLLKIGKRNFYIIKWLQL  
477 aa / 34eme aa / 0,9785

>*Candida.glabrata*-mito

MIRCRMILNRKGRFLVSNFRRSYNVLRRELDRGLVSVQVSAPEEKLEEQLESGQKVKLYCG  
VDPTAKSIHLGNLVPLMVLHLYFVRGHDIVTLVGGATGCVGDPGSRRTTTERSLMEDKTRLS  
NVSSISGQLEKFFINGLAYYNTKNKDRNETLPGKHIKENNYNWWKDVKMLDFLANYGRHI  
RVQSM LARDSVTARLNTQNSLGFNEFTYQILQAYDFHNLNKTYGVS IQVGGNDQWGNITA  
GIDLISRLSSRTKQAYGVTVPLLQTASGEKFGKSAGNAVFIDPSISTPYDIFQFFLNTLD  
EDVPRFLKTFSTLIEEIESMTKRHNEKPHRLRGQKMLAKEITDMLFGVNGDKAQSISN  
IIFGTMEDKEMKLSGLELIDMFREARILQOAKRGESLVALVSRLTHSSKSEARRKIEQGS  
IYLHTRS VKITKDINDYNDYLIDDKVLLLRIGKQKCFVIEMQ  
462 aa / 23eme aa / 0,9882

>*Ciona.intestinalis*-mito

MSTPVLVHTFTTKSWIRCSKLVTAFRGKFLPTPNIKYTTAPSRTYSTNYETNKT KIRPKN  
VLK LHERGILQDIYPPAAVGLPGLLESKQOCFYCGFDPTADGLHVGNNLSIMVLLHCQRA  
GHNVI AVIGNATVRAGDPSEHTKDRNMLNDNAIQDNVTAIQDCIAKIVDTHANYIYDNKH  
RKKLPDFKILFNQ TWYKELNAIDFISTACRNFRIGPMMEKRYIKDRIQSSEGLMLAEFMY  
QVLQAYDWQYLHDNYNCTFQVGGVDQLGNINAGQQLIRKTNKGQVYALLTPLVTNLRGQK  
LGKTAGNSVWLNANKTPPF EFYQFFFQPD SMVETFLKYFTFIDLKRINEIMIEHTKHPH  
LRNAQRILAKHVCKLVHGPEGLESAERCTRALFGGSLEDIEKLSEEEIMTTFNATNIVTA  
VLEPDTTVYDVATHAKAIP LGNKGENLIRAGGVRINGETVDSPHLILMNGEHVLKNDLTV  
LSVGKTNHFIIKWKLPRIKHDEDSVRETEYPVPM  
514 aa / 46eme aa / 0,9716

>Cavia.porcellus-mito

MAAPVLRFLFPGWWSGTSGLLGLLPRGVRQSHSTAQGFLLAAQKARGLFKEFFPETGKTKE  
LELFDRRSAGNSPQTIYCGFDPTADSLHVGHLALXXXXXXXXXXXXXXXXXXXXXXXXXXXX  
XX  
XX  
XXXXXXXXXXXXXXXXXXXXXXXXXXLTGEDVFGITVPLITNTTGAKLGKSAGNAVWLDKRVKTS  
PFELYQFFVRQPDDSVRYLKLFTFLPFAEIDHIMQLHGKEPEKRGPKRLAAEVTKLHVGQ  
EGLDSAKRCTHALYHSSIDALEVMSDQELKELFKEASFSELVLDPGTSLVLDTCRKANAI  
PDGPRGYRMITEGGVVSINHRQVTNPESVLVVGQHILKNGLSLLKIGKRNFYIIKWLQL  
477 aa / manque 165 aa

>Ciona.savignyi-mito

MANRTFVRLFTATSRTNITTPYKLNLTKSACRIARYYSASTSTRHYSANNSTRHYSENN  
STRHYSANNSTRHYSANNSTRHYSANDSTKSKTDIRPKNVLNIHERGILEDIYPPAAVGL  
PQLLESKQOCFYCGFDPTADGLHVGNLLSIMVLLHCQRAGHKAIAVIGNATVRAGDPSEH  
TKDRKVLSDAEIQMNVEAIRDCITTI VGNHAQHIYDNKHKKNLPEFKVLYNQTWYKKLNA  
IDFISSACRNFRIGSMLEKRYIRDRINSSDGLMLAEFMYQVLQAYDWKYLHDNHCNFQV  
GGADQLGNIHAGQQLIRKSSKKKVYALLTPLVTNLRGQKLGKTAGNSVWLNPKKTAPFEF  
YQFFYKQPDMSVETFLKYFTFIDMKMINEIMIDHEKQSNLRQAQRILAKHVCKLVHGPEG  
LKSAERCTRALFGGKLEDIEQLTEDEIRSTFDATNVINAVLEPDTTVYDIAAFKAI  
PPGNKGENLIRAGGVRINGETVDSPHQILMSGAHVLKNDLTVLSVGKTTHYIVKWKLPRVKHD  
EDSVLEMEYPMKT  
553 aa / 84eme aa / 0,9930

>Debaryomyces.hansenii-mito

MISKLISGNRARTLFPFVSSLTSNKRFQTTMPAILKEDLPETVTI I PTISELTKLGDFS  
NDIDVSLVGHLSRYLVESITHDDLKFWTDPSSSEKRFKLYCGADPTAESLHLGNLLPLMV  
LLHFNLRGHDVVGLVGGATGAVGDPGSRGRTTERTQLAEKRLDNVNR IQNQLSNFLKQGVE  
YAKSRNFP IQKQGI FTENNAKWWESI KILDFLSTYGKHIRVSAMLRDS IQSRLKSEHG  
LGFNEFSYQVLQAYDFWHLFKNEKVNLI GGNQWGNITAGVDLISRLQKHFNRVSKPKD  
KIEEPSFGMTVPLLTPTPTGEKFGKSAGNAIFINEKLTTPYQLYQYFINTPDEIVAKLLKV  
FTILPMATIEKIVQKHNEPDIRAAQRILAREVTDLIHGNGMGDKMAYVTSFLFPPTDQF  
FIDDISANKLISVFEKSGILVRFKFDDFNIDDLKSTLLAKVMKSKTETKNLIAAGGIY  
LGLDRNQHIDPEDVVLFDKDHLIDGKLLLVVRAGKQNYVVVEFS  
523 aa / 29eme aa / 0,6716

>Drosophila.melanogaster-mito\_SYYM\_DROME

MLPLRRSLLKPLQDVLRRHSRQMSQKNLLELTDGRGFFHGIFPDTAAPRMKQLFTRGQOSI  
YAGFDPTADSLHVGNNLVIMGLIHCQRAGHRPIALVGGATGLIGDPSGRKTERNQLGETV  
IETNLKAI EQQLRRVFENHENCLWDSKKQKLPLAPLI IVNNADWYADLQLIDFVANMGRH  
FRMGSMLSRSSVQSRLSEEDGMSFTEFTYQIFQAYDWLHLLRRHNCCFQMGGSQDTGNLM  
TGHELISRVERKREVFGLTLPVTTEEGDKFGKSAGNAVWLDGNKTS PFALYQFFLRMPD  
SEVEKLLKLF TFIPLPQVEQLMREHTKEPEKRAQTLAEDVTLVHGESGLKQAERTN  
ALYKGNVEGLAELNLSEIQOTFQATMVNLLTEPGMSILELAMKAKCFPTETDAVRI  
INAGGFYVNQKRVQNI AEVLT TG VHI LRNGISLLRVGKRNFYIVRWQ  
464 aa / 23eme aa / 0,9298

>*Drosophila.pseudoobscura*-mito

MLQLRWSLLRPLQGVCNRPRRYFAHKNLLELTDRGFFQGI FPD TAAPKIKQLFTSGQQSI  
YAGFDPTADSLHVGNNLLVIMGLLHCQRAGHRPIALVGGATGLIGDPSGRKTERNQLGESI  
IESNLKAI EQQLTRVFENHEKCLWDNRKNSTPLQPLTIVNNAEWYANLHLIDFIANMGRH  
FRMGSMLSRSSVQSRLESEDGMSFTEFTYQIFQAYDWLHLLRKHNCCFQMGGS DQTGNLM  
TGHELISRVERKREVFGMTLP IVTNEEGDKFGKSAGNAVWLDENKTS PFALYQFFLRMPD  
SEAEKLLKLFTFIPLTEIEQLMQEHEKEPEKRKAQTL LAEDVTLLVHGENGLKQAERTN  
ALYK GKVDGLAELNYAEIKQTFQGA AVVDLLTEPGMSILQLAMKAKCFPTETDAVRIINA  
GGFYVNQKR VQNI AEVITAGIHILRNGVSLLRV GKRN FYIVRWQ  
464 aa / 31eme aa / 0,9889

>*Danio.rerio*-mito

MAASIARSCCRVKSHFILRKTSYCKLLFHSSASKTSSLLSSLHNRGLLKDSFPEVAAQAE  
IPDLLRSGPQSIYCGFDPTADSLHAGNLLAII GL LHFRSAGHHI IALLGGATAQIGDPSG  
RQTERERLSPA AVQENARGILESLHRI FTHHELYFCPDS SRLGRLTVLNNRSWYKDYDLL  
EFFSEVGRSFRMG TMLSRHSVQTRLKSAEGMSFTEFSYQLFQAFDFYQLHQLHGCR IQLG  
GTDQLGNIMSGHEFIHRVGGFEEVYGVAVPLVTTSMGDKLGKTAGNAVWLN RDKTS PFEF  
YQYFLRLPDNSVERYLKLFTFLSLSEVENLMEQQRKDPGKRI AHKRLAAEVTKLVHGKDG  
LES AKRCTTALYHSSIEALEQMSDTELQEI FREAPFHEFFLEPGTTVLDACRRARAI PDG  
PRGYHMIK DGGVWINHQRAENPEQVLI VGGQILSNGLSLIRV GKKNFHILKWL SL  
475 aa / 21eme aa / 0,9951

>*Felis.catus*-mito

MAAPMFRFCFSRGRWFGSPGMSGVLFGLGLREAHSGAQGLLAVQKARGLFKEFFPERGTKIE  
LPELFDRGTGNFPQTIYCGFDPTADSLHVGHL LALLGLFHFQRAGHNVI ALVGGATARLG  
DPSGR TKEREALGAERVRSNARALRQGLEAMGANHQQLFADGRPWGSFTVLDNSAWYQKQ  
HLVDFLSAVGGHFRMG TLLSRLSVQTRLRSPEGMSLAEFFYQVLQAYDFYYLFQHYGCRV  
QLGGSDQLGNIMSGYEF IHKLTGEDVFGITVPLITSTTGAKLGKSAGNAVWLN RDKTS PF  
ELYQFFVRQDDLVERYLKLFTFLPLPEIDHIMQLHEKEPEKRGPQKRLAAEVTKLVHGQ  
EGLSAKRCTQALYHSSIDALEVMSDQELKELFKEASFSELVLDPGTSVLDTCR KANAIP  
DGPRGYRMITEGGVSINHRQVTNPESVLVVGQHILKNGLSLLKIGKRN FYIIKWLQL  
477 aa / 23eme aa / 0,8672

>*Gasterosteus.aculatus*-mito

MAASMA RTCRHVPRRGPCVLRRLTLRSKLHNGLLLSLNKR GILKESFPEDAAQDRLPRL LQ  
SGAQTVYCGFDPTADSLHVGNNLLAIVGL LHFRSAGHHVLAVLGGATAQIGDPSGKTSERE  
RLSAEVAEANTRGIRESIQRVFTNHELHFHDGSRKAGTVTVLNNLSWYKNRGVVDLFLSEA  
GRHLRMGTMLSRHSVQSRLRSADGMSL TEFTYQVFQAYDFHHLNQIYGCKIQLGGTDQLG  
NLMSGHEYIHKASGDEVFGLTVPLVTSSAGDKLGKTAGNAVWLN RDKTS PF EFYQFFLRQ  
PDASVERFLKLFTFLPLAEVERLMEQQRDDPGKRLAHKRLAAEVTKLVHGKEGLESAKRC  
TDALYHSSVQALEEMSDGELQELFREAPFHELLLEPGTTVMDACRRLSAVPEGPRGYQMV  
SEGAVWINHRRTDNPEQLLLPKLHILSNGLTLLRVGKRN FYIIKWL SLQGLS  
472 aa / 16eme aa / 0,9951

>Gallus.gallus-mito

MAAPTLRRCRAAGLWALRRARPPGTRHVHERPQRARGAKGLLAEQCERGLFQEVFPA  
QGADEQLEALLEPGRPPVAVYCGFDPTADSLHVGHLPAVMALLHFQORAGHTVLA VVGAT  
ARLGDPSGRERAREPLEAERVRAHARGLREGLGRLLNHRALFWPAGGGRPLGRAELLDN  
AWWLSREPLLDLFCGAGGRLRMGTLLSRQACQARLRS AEGMSLAEFLYPALQAYDFLHLH  
RHHGCRVQLGGADQMGNIMSGYELVTKMTGTDVFGITVPLITSTTGDKLGKTAGNAVWLN  
RDKTSPFELYQFFVRQODDIVEKYLKLF TFLPLEEIDHIMEMHAREPEKWGPQKRLAAEV  
TKLVHGREGLSAKRCTKALYHSSVEALEAMSDQELQELFRQAPSAELILEPGTTLLDLC  
RKANAIPHGPSGYQKITDGGVSVNGIRVTNPETVLI LGQHILKNGVSLLRIGKKNYI I K  
WLQL

484 aa / 39eme aa / 0,9971

>H.sapiens-mito

MAAPILRSFSWGRWSGTLNLSVLLPLGLRKAHSGAQGLLAAQKARGLFKDFFPETGTKIE  
LPELFDRGTASFQTIYCGFDPTADSLHVGHL LALLGLFHLQORAGHNVI ALVGGATARLG  
DPSGRTKEREALETERVRANARALRLGLEALAAHQQLFTDGRSWG SFTVLDNSAWYQKQ  
HLVDFLAAVGGHFRMG TLLSRQSVQLRLKSPEGMSLAEFFYQVLQAYDFY YLFQRYGCRV  
QLGGSDQLGNIMSGYEFINKLTGEDVFGITVPLITSTTGAKLGKSAGNAVWLN RDKTSPF  
ELYQFFVRQPDDSVERYLKLFTFLPLPEIDHIMQLHVKEPERRGPQKRLAAEVTKLVHGR  
EGLDSAKRCTQALYHSSIDALEVMSDQELKELFKEAPFSEFFLDPGT SVLDTCRKANAI P  
DGPRGYRMITEGGV SINHQV TNPESV LIVGQHILKNGLSLLKIGKRNFYI IKWLQL

477 aa / 32eme aa / 0,9740

>Kluyberomyces.lactis-mito

MIFGRLLVRNLATKCDVLQVLRERGLVQQVSQPEKLLSEKLSNGDKIKLYCGADPTAKSL  
HLGNLLPLMVLLNFYVRGHDVVSLIGGATGKVGDP SGRKTERDVIAEAKRQGNIERIVEQ  
YKTFFTNGLKYYESRIKERDAPGKVTYLNNISWWQDVKMLDFLATY GKHIRIQNMLSRDS  
VSSRLQSQDGIGFNEFTYQILQAYDFYHLYSNEQVSIQVGGNDQWGNITAGIDLINRISP  
QHVKTRPAFGLTVPLLTSTSTGEKFGKSAGNAVFIDPEINTSYDMYQFFVNTTDADVAKQL  
KIFTLLPLELIEDIMAQHALSPNERYAQKRLAKEVVDLIHGIGKGD DAEFVSRILFGGSG  
FENVKAAELIRVFDENRILNKIPINTPLSDIVCQLIDCSKTESRRRIKQGS IYLGPNKNK  
VDDTTNFSPFLIDERVLLLRVKGKQKCYVIECI

453 aa / 19eme aa / 0,9064

>Monodelphis.domestica-mito

MAAPMFQVFRRLFGLGRWPGPRLSGRRCKGHFGARDVLLAQKMRGLFADIFPAKGVAPEL  
PELLAGGAQTVYCGFDPTASSLHVGHLLPVLG LLLHFQORAGHHVIAVVG GATARLGDPSGR  
SKERTALSAERVRD NAKALRASLEQLAAEHRRLFHDGRPWGTF TFLDNAAWYSHLALVDF  
LGAVGGSFRMGPLLSRHSVQARLGSPEGMSLAEFLYPVLQAYDFYHLHQRYGCRVQLGGS  
DQLGNIMSGHEFVRRVTGKDVFGITVPLITSTAGAKLGKSAGNAVW LHRELTPFEFYQFF  
VRQPDDCVERYLKLFTFLPLPEIEHIMQM HMAPEKRGPPQIRLATEVTKFVHGQEGLGSA  
KRCTQALYHCSIDALEIMSDEELKELFKEAPFSELVLEPGT SVLDTCRKANAI PDGPRGY  
QMITEGGV SINHKPVTNPESV LVVGEHILKNGLSLLKIGKRNFYVIKWLQL

471 aa / 29eme aa / 0,9944

>Macaca.mulatta-mito

MAAPIFRSFSWGRWSGTLNLSVFLPLGLRKAHSGPQGLLAAQKARGLFKNFFPETGTKIE  
LPELFDRTGASFPQTIYCGFDPTADSLHVGHLALLLGLFHLQRAGHNVIALLVGGATARLG  
DPSGRTKEREALETERVRANARALRLGLEALAAHQQLFTDGRSWGSTVLDNSAWYQKQ  
HLVDFLAAVGGHFRMGTTLSRQSVQLRLKSAEGMSLAEFFYQVLQAYDFYYLFQRYGCRV  
QLGGSDQLGNIMSGYEFINKLTGEDVFGITVPLITSTTGAKLGKSAGNAVWLNDRDKTSPF  
ELYQFFVRQPDDSVR

316 aa / 17eme aa / 0,9546 / tronquée en C-terminal

>Mus.musculus-mito\_SYYM\_MOUSE

MAAPMLRRLCRVPQSLVWLGRSRAVRPGARGMLVAPRARGLFKEFFPESGKTELPELFD  
RRRAGSSPQTVYCGFDPTGDSLHVGHLTLLGLFHFQRAGHNVIALLVGGSTALLGDPSGR  
TKGREALSAECVRANAHALRRGLEALAAHARLFADGRPWGSFTVLDNAAWFQEQHLVDF  
LATVGGHFRMGTTLSRSLVQSRLKSPGMSLAEFFYQVLQAYDFYYLFQHYGCRVQLGGS  
DQLGNIMSGYEFIHKLTDGEDVFGITVPLITSTTGAKLGKSAGNAVWLNREKTS PFELYQF  
FVRQDDSVRYLKLFTFLPLPEIDHIMQLHVKEPEKRVAKRLAAEVTKLVHGQEGLD  
AKRCTQALYHSSIEALEVMSDQELKELFKEASFSELVLDPGTSVIDTCRKANAI PDGPRG  
YRMITEGGV SINHRQVTNPESVLVIGQHILKNGLSLLKIGKRNFYIIKWLQL

472 aa / 46eme aa / 0,9995

>Neurospora.crassa-mito\_SYYM\_NEUCR

MLLRRTKALIRSGGSIKAYAAANPSCFILQRRGLRREFGPKYTAKINEAEENWQARAEAIK  
KGKKQNTWDLFEERGYVKDTAGTKEHIAELMRTRRIGAYVVIDPTAPSLHVGHLPLMPL  
FWMYLEGYKAFTLIGGSTAKIGDPTGRLKSRDHLSSSDATMNMTKIHYQLKKLWENVDTQ  
MRARGYEADWARKRGI VNNNNHWNKQPMLEVLRRVGHALRIGPMLSRD TVKNKMTQGDGV  
SFAEFTYPI MQGWDW FELFYQQGVQM QIGGSDQYGNII SGLEVVKAARESE PDPQERKYV  
TPKTALDECVGFTVPLLT DSSGAKFGKSAGNAIWLDPYQTSVFDYGYFVRRSDQEVENTL  
LKLFTFMP ISEITKTMEEH IKDPSKRVAQHTLAREVVTLVHGKQEA SAAEDQHRMMYTGQ  
MTI PQVSRAKDAATGGDQYKTI SDQPVT LNNAPRIDMILPESLIMGKSIGRILYAAGLAS  
STTEGHKLA A AQGCYVGGAH RAGGENVTMNPDLISFMPVKLWFPGETQRYLINGNLLILR  
KGKHNVRVIQMVSDVEYAASGQTPGQSFTGAVRKLNEIMKNLKEKKTPEEAKNAVNEL  
QKSSQEKQOQQOIIIFPEEKSRQKKDMETK LKQEMIASVKTIDGMMDEKPSVRGDGVKKQT  
QDDRDPYKW

669 aa / 45eme aa / 0,9895

>Oryctolagus.cuniculis

MAAPMLRCCSRARWSGAPSPSRVLRSGLRGAHSGPQGV LAAQKARGLFKEFFPERGKTE  
LPELFDRTGSGSFSFPQTVYCGFDPTSDSLHVGHLALLLGLFHFQRAGHNVIALLVGGATARL  
GDPSGRTKEREALGAECVGANARALRAGLEALAAHQQLFGDGRAWGSFTVLDNSAWYQG  
QHLVDFLAAVGGHFRMGTTLSRSLVQTRLKSPGMSLAEFFLYQVLQAYDFYYLFQRYGCR  
VQLGGSDQLGNITSGYEFVHXX  
XXXXXXXXXXXXXXXXXXXXYLKLTFLPLPEIDHLMQLHLKEPEKRVQKRLAAEVTKLVHG  
REGLDSAKRCTQALYHSSIDALEVLSQELKELFREASFSELVLDXXXXXXXXXXXXXXXXXXXX  
XXXXXXYRMITEGGV SINHRQVTNPESVLVVGQHILKNGLSLLKIGKRNFYVIKWLQL

478 / 15e aa / 0,9936 / manque 57 et 21 aa

>Oryzias.latipes-mito

MAAAVVRVCERGLLRSSALRAFGSPLRRSSSAGGGLLLSLHRRGVLKDSFPENAAQDRL  
PQLLQSTPQTVYCGFDPTADSLHVGNNLAI IGLLHFRGAGHHVIAVLGGATAQIGDPSGK  
TSERERLSVTSVEANTQTIRESIQRIFTNHEMYFHDSGRTLGTATVLNMSWYKSWGVE  
FLSEAGRHLRGMGTLRHSVQARLKS PDGMSL TEFTYQVFQAFDFHHLNQTRGCR IQLGG  
TDQLGNLMSGHEYIQKVS GEDVYGLT I PLVTNSMGDKLGKTAGNAVWLNDRDKTSPFELYQ  
FFLRQPDANVEGYLKLFTFLPLAEVESL MKLQREDPGKRLAHKRLAAEVTKL VHGKEGLE  
SAKRCTNALYHSSVRALEEMSDEELQELFREAPFHELLLEPGISVIDACRLVSAIPEGPR  
GYQMVS DGA VWINHTRADKPEQVLV PKLHVLSNGLTLLRVGKRNYYI I KWLGL  
473 aa / 45eme aa / 0,9943

>Oryza.sativa-mito

MASIAMAASLRTFLRPHRCRLLLRQSRSL SASAAPAGAAA ASTVIRRSVVEVLRERGLVE  
ATTSESLGSASASPRELKAYCGFDPTAESLHLGNLLGLVVL SWFRRCGHNAVALVGGATG  
RVGDPSGKSAERPELDLAAVETNSNAIKSLIGQILDRAPEPSQHSQSGKNLNLEQNEQAL  
VNSGEKMG SFQILDNYDWWKDITLLDFLKEVGRFARVGTMIAKESVKKRLMSEDGMSYTE  
FTYQLLQGYDFLYMFRNMGVNVQIGGSDQWGNITAGTELIRKILQVEGAHGLTFPLLLKS  
DGTKFGKTEDGAIWLSPKMLSPYKFYQYFFSVPDIDVIRFMKILTFMSLDEIQELEESMK  
KPGYVPNTIQRRLAAEVTRFVHGEEGLEEALKATEALRPGAQTQLDSQTI EGIADDVPSC  
FLPYGQVLKSPVLVDLAVSTGLLASKSAVRRLIKQGGLYLN NIRIDSEDKLVEEGDVVDGK  
VLLLSAGKKNKMVVRIS  
497 aa / 49eme aa / 0,9977

>Podospora.anserina-mito

MSMSRGSVCRRCLLTMKSMAGGGPTSTY AQQRGKKTWHGPKYQAKIDQAQADWEERA EKI  
KKGEIQHTWDMFVERGYVKDTAGSHETIRKMLHKRIGAYTGIDPTAPSLHIGHLLPLMP  
IFWYMHGYAGYTLIGGATAKIGDPTDRLVSRTP LKRTDLTMNLTKIHYQLKALW MNVEE  
QARRRGFEKDWAWKRAVVNNSTWWSLPLIEVLKRLGDSMRMGPLLSRDTVKNKMSKGDG  
MSFSEFTYPLMQGDWWHMYQANGIQMQIGGSDQYGNIVTGVETVKVVRDNEPDPAKKIE  
GGPFNDPVGFTVPLLTDSAGVKFGKSAGNAFLDKFQTSEFDLYGYFVRRSDQEV EKLKLL  
FTFLPMENINEAMKIHSEN PARRVAQHLLAFEV VGLVHGMNAAHRTALNHQARYGKQIDI  
PGVTLRMPKAATEDTPPSILDAPKMDMQLPESLIMGKSI GRILYAAGLAKSASEGHRLAT  
QQGAYIGAMPGHKRTEDNKVMDYSQLSFTPIK LWFQETRNYLIDGKLLILRKGVQIRV  
IEMVSDEEWKESGQTYPGEPGTGALRMLRQQLKMLKSGMLTPDEVKANLKNHVEEEAPP  
GFMKFPDQDSYAIRRATQELMDEIHQKEVGGDSPREERRE  
640 aa / ? / 0,8646

>Populus.trichocarpa-mito

RSSVVDILEERGLLESITSDNLRISITSTL KAYCGFDPTAESLHLGNLLGIIVLSWFQR  
CGHKAVALIGGATARIGDPSGKSLERPELDADTLENNTQGITNVITRILNMNSSSNVNGD  
GNHLNSSSFVVMNNYDWWKEVRLLDFLKQVGRFARVGTMMGKESVKKRLESEQGMSYTE  
FTYQLLQGYDFLYLYQNEGVNVQIGGSDQWGNITAGTELIRKILQPEGDVAFGLTFPLLL  
KSDGTKFGKSEDGAIWLSPLLSPYKFYQYFFSVPDADVIRFLKILTFLDIEEIDELEKE  
MNRPGYTPNTAQRRLAEQVTLFVHGEGDGLNEALKATEALRPGAETKLDWKT FEGIAEDVP  
SCSLASDQVLNISLIDLSVSSGLLDSKSAARRLLKQGGLYLNNSRV DSETKRIEQDQDIVD  
GKVLLLSAGKKNKVIVRIT  
tronquée en N-terminal



>Pan.troglyodytes-mito

MAAPILRSFSWGRWSGTLNLSVLLPLGLRKAHSGAQGLLAAQKARGLFKDFFPETGTKIE  
LPELFDRTGASFPQTIYCGFDPTADSLHVGHLLALLGLFHLQRAGHNVIALVGGATARLG  
DPSGRTKEREALETERVRANARALRLGLEALAAHQQLFTDGRSWGSTVLDNSAWYQKQ  
RLVDFLAAVGGHFRMGTTLSRQSVQLRLKSPEGMSLAEFFFYQVLQAYDFYYLFRYGRV  
QLGGSDQLGNIMSGYEFINKLTGEDVFGITVPLITSTTGAKLGKSAGNAVWLNDRDKTSPF  
ELYQFFVRQPDDSVRYLKLFTFLPLPEIDHIMQLHVKEPERRGPQKRLAAEVTCLVHGR  
EGLDSAKRCTQALYHSSIDALEVMSDQELKELFKEAPFSEFFLDPGTSLDTCRKANAI P  
DGPRGYRMITEGGVSINHQQVTNPESVLIVGQHILKNGLSLLKIGKRNFYIIKWLQL  
477 aa / 17eme aa / 0,9740

>Ratus.norvegicus-mito

MAAPMLRHLCRVPQSGVWTRGPRAVRPGARGMLVAPRARGLFKEFFFPESGKTTELPELFD  
RRRAGSPQTVYCGFDPTGDSLHVGHLLTLLGLFHFQRAGHNVIALVGGSTALLGDPGRT  
KEREALSAECVRANARALQRGLETLAANHARLFADGRPWGTFVLDNAAWFQKQHLMDFL  
ATVGGHFRMGTTLSRSLSVQSRKSPGMSLAEFFFYQVLQAYDFYYLFRHYGCRVQLGGSD  
QLGNIMSGYEFIHKLKLTGEDVFGITVPLITSTTGAKLGKSAGNAVWLNREKTS PFELYQFF  
IRQDDSVRYLKLFTFLPLPEIDHIMQLHVKEPEKRIAQKRLAAEVTCLVHGQEGLSA  
KRCTQALYHSSIEALEVMSDQELKELFKEASFSELVDPGTSVIDTCRKANAI PDGPRGY  
RMITEGGVSINHQRVTNPESVLIVGQHILKNGLSLLKIGKRNFYIIKWLQL  
471 aa / 46eme aa / 0,9980

>Saccharomyces.bayanus-mito

MLELRNCSHLVSSGKRIVPLVTRHRINTKVWHKTRLYSCTASPSETQDTPNSWNSNLVDE  
LKQRGLVCQVSQPEDDLRMLNSNKKIKLYCGVDPTAQSLHLGNLVPLMVLLHFYVKGHD  
VVTVIGGATGKVGDPGRKTERDVMENDIRQNNVQSI SAQLORFFQNGLYYTD RDRDKRG  
IESYGKYTPRNNFNWWDIKMLDFLADFGKHIRVQSMLARDSVSSRLESQNSLGFNEFTY  
QILQAYDFYHLYKEENVSVQVGGNDQWGNITAGIDLINRLOPTQKKGVPFGVTVPLTTA  
TGEKFGKSAGNAVFIDPSINTAYDIYQFFFNTLDADVSKFLKIFTFLLSSNEINSVMEMHI  
KSPNLRGQMLLAKEATDMLYGVGSGTDSETLSNVI FGHYDGTLSAAKLKELCHKAKILQ  
HADKKNLVLKLSKLGDCSISEAKRKL SQSIIYLHHSRIKVKEDVAEWGPYLIDDEVLI L  
RMGKQKCFI IEMH  
493 aa / 38eme aa / 0,9609

>Saccharomyces.castellii-mito

MLKWGVARRWITSTACRSNISELKQRGLVSQISQPESWLSRSLKEGKKIKLYCGVDPTA  
KSLHLGNLVPLMVLLNFYVRGHDI VTLVGGATGRVGDPSGRKSER SAMEDLVRETNVNSI  
GKQLQRFFLKGKEYYETKIRENESRRFGEHILEDNYHWWKDVKMLDFLAQYGRHIKI QSM  
LSRDSVAARLSNQNSMGFNEFTYQILQAFDFYHLYKEHGMVQVGGNDQWGNITAGIDLI  
GRLEEKIKEKPAFGITVPLTTSTGEKFGKSAGNAVFIDPEVNTPFDIYQFFYNTTDADV  
ERFLNIFLLTSSEITEIVKEHAKNTHLRSGQKILAREITELLHGKDSSEAEQVSDILF  
ASVRSGLSLSGETLVNICKSANILOYGSKSETLIELIARLTNGSKSEAKRKIAQGSVYL  
HSDRIKAIDNVTNWEPYLIDKTVLLMKLGKQNTFVVKLQ  
459 aa / 37eme aa / 0,9329

>Saccharomyces.cerevisiae-mito\_SYYM\_YEAST

MLELRSCSNLVNSSRRLVPLVTYSGLSAITLPKSRFYQSPPSALEVQGTSDSRSDNIDDEL  
KQRGLVSQVSQPESFLRRTLKLNNDKIKLYCGVDPTAQSLHLGNLVPLMVLLHFYVKGHDI  
VTVIGGATGKVGDPGRKTERDVMENDIRQSNVASISQQLQRFFKNGLEYRNRICALTED  
VPSGKYTPRNNFNWWDIKMLDFLADFGRRHIRVQSMLARDSISSRLQTKNGLGFNEFTYQ  
VLQAYDFYHLYKEENVTIQVGGNDQWGNITAGIDLINRIQPIKNKGLPFGITVPLLTTAT  
GEKFGKSAGNAVFIDPSINTAYDVYQFFYNTLDADVPKFLKIFTFNLNSSEIKKIVETHIK  
SPSLRYGQTLAKEVTDMLYGVGSGSDSEALSNIIFGRYDGTLSAAKLVDLCKKARILQY  
ADREIDLICKLICKLVNCSVSEARRKLSQGSVYLLHSSKSKVNNENISNLAPFLIDDRVLILR  
IGKQKCFIEMR

492 aa / 38eme aa / 0,8476

>Saccharomyces.kudriavzevii-mito

MLEIRTFSHVLNSTKIKRLVTDYGLSAALWRKARFYSHTVGPSEPQDTSNSRSDNIDDEL  
KQRGLVCHVSQPENVLWTRLKGDDEKIKLYCGVDPTAQSLHLGNLVPLMVLLHFYVKGHDI  
VTVIGGATGKVGDPGRKSERDLMKNDVRQSNVKSISQQLRRFFQHGLEYYNMRHNVTEF  
SSPGKYTPRNNFDWWDIKMLDFLADFGRRHIRVQSMLARDSVSSRLQTQNSLGFNEFTYQ  
ILQAYDFYLYKEENVRVQVGGNDQWGNITAGIDLINRLQPTQKNGLPFGVTVPLLTTAT  
GEKFGKSAGNAVFIDPSVNTAYDVYQFFNTLDADVVKFLNIFTFLDPNVITSVVESHNK  
SPNLRYGQTVLAKEVTDMLYGVGSGSDSETLSNVIIFGRYDVTLSATKLIELCKKAKILQS  
ADRKDSLVLICKLADCSVSEARRKLSQGSVYLLHSSRIRIKEDVSNWAPYLIDDQVLILR  
IGKQKCFIIMR

492 aa / 37eme aa / 0,9161

>Saccharomyces.mikatae-mito

MLELRSCRYLLNSSKRFVQPVTCHYLRAIACHKVRFYQTVSPSEVRETLNSQSENILDE  
LKQRGLVCQVSQPEVALQTRLNSDDKIKLYCGVDPSAQSLHLGNLVPLMVLLHFYVKGHD  
IVTVIGGATGKVGDPGRKTERDVMKNEIRQNNVTSISQQLQRFFRNGLEYRNRALYALTR  
DVSSGEYTPRNNFNWWDIKMLDFLADFGGHIRVQSMLARDSVSSRLQTKNSLGFNEFTY  
QILQAYDFYHLYKEENVTIQVGGNDQWGNITAGIDLINRMQPIQKNGLPFGVTVPLLTTA  
TGEKFGKSAGNAVFIDPKVNTAYDIYQFFNTLDADVPKFLKMF'TFLDSNEINSVVESHG  
KSPNLRHGQAFLAKEVTDMLYGVGSGSESEALSNIIFGRYDETLASASKLIELCEKARILQ  
HANRKDDLICKLICKLIDCSVSEARRKISQGSVYLLHYSRTKVNENISNLAPFLIDDRVLIM  
RIGKQRCFIIEMP

493 aa / 38eme aa / 0,9012

>Schizosaccharomyces.pombe-mito

MSRLLACLKQLQARSLIHNTTLLQPCNVNSVYLGADPTAASLHVGNLVALMPLVHFFLN  
GFPVFTVIGDATAQLGDPGRSTSRKQMAETTRTANSNSIHNQLKDLSSSILSYAQDCNY  
PFSQMPSSSQWSIVRNSSWYENLKLKFLSSVGPVHRVVSQMLARDSVTTRLQSPSGLSFA  
ELTYQLLQAYDYSYLYENHSVNLQIGGSDQWGNITAGTDLVRRTHPNANVYAL'TPPLLTS  
SSGQKLGKSAGNAIWLDPKLTDSYSLYQYFISAPDDLACKCLDMLTLLPLEQLEQIKAEH  
EKDPSQRIVHKYLASNVVRMVHGKKALELAQIQTKLLHGAHQAPFGFYSEAPQQGDSFPS  
LPEIRALFKDCKFYRTIDSSIKDQPPSRLLRTLQIYTSRKEATEHILSGAVSLGHKPIILD  
SNYKFPDNLFLVLRAGKRTFVLDSL

445 aa / 13eme aa / 0,7418

>Sus.scrofa-mito

MAAPMLRRVSRGRWLGTTPGQSGVLPGLRKAHSGAQGLLAVQKARGLFKEFFPERGTKTE  
LPELFDRLAGNVPQTIYCGFDPTADSLHVGHLALLGLFHFQORAGHNVIALVGGATARL  
GDPSGRTKEREALDAERVQSNALALRQGLQALAAHQQLFTNGRSWGSFTVLDNSAWYQK  
QHLVDFLATVGGHFRMGTLTLLSRLSVQSRLLKSPGMSLAEFFLYQVLAQYDFYYLFQHYGCR  
VQLGGADQLGNIMSGYEFIHKVTGEDVFGISVPLITSTTGAKLGKSAGNAVWLNREKTSP  
FELYQFFVRLQDNLVERYLKLFTFLPLPEIDHIMQLHVREPEKRGPKRLAAEVTKLVHG  
QEGLDSAKRCTQALYHSSIDALEVMSDQELKELFKEASFSELVLDPGTSLVLDTCRKANAI  
PDGPRGYRMITTEGGVSIHRQVTNPESVLVVGQHILKNGLSLLKIGKRNFYIIKWLQL  
478 aa / 23eme aa / 0,9886

>Tupaia.belangeri-mito

MAAPMLRCFCRGSWSGPRGPGSVLRPGLRKAHSGAQGLLAAQRARGLFKEFFPERGTKIE  
LPELFHPGTAGYSPQTIXX  
XX  
QHLVNFLAAVGGHFRMGTLTLLSRLSVQARLKSPEGMSLAEFFYQVLAQYDFYYLFQRYGCR  
VQLGGSDQLGNIMSGYEFIHKLITGEDVFGITVPLITSTTGAKLGKSAGNAVWLNDRDKTSP  
FELYQFFVRQDDCVERYLKLFTFLPLPEIDHIMQLHIKEPEKRAPQKRLAAEVTKLVHG  
REGLSAKRCTQALYHSSIDALELMSDQELKELFKEASFSELVLDPGTSLVLDTCRKANAI  
PDGPRGYRMITTEGGVSIHRQVTNPESVLVVGQHILKNGLSLLKIGKRNFYIIKWLQL  
478 aa / 45eme aa / 0,9919

>Tetradon.nigroviridis-mito

MAASMVRTFYRALRHRTCVLRSRPTCFNRMLQYEFHSSSTLLNGLLFSLYKRGVLKESFPE  
SAAQDQLPQLLQAGSQTVCYCGFDPTADSLHVGNNLALLIALHFRNAGHHVLAVVGGATAQ  
IGDPSGKARERERLSADTAEANTRGIRESLQRIFTNHELIFHDSAKKLGTVTVLNNLSWY  
KDRGVVEFLAEAGRHRMGTMLSRHSVQARLRSPDGMSLAEFFTYQVFQAHDFYHLNQIYG  
CKIQLGGTDQLGNLMSGHEYIHKVSGEVYGLAIPLVTSVGDKLGKTAGNAVWLNDRDRT  
SPFQLYQFFLRLPDSGIERYLKLFTFLPLAEVERLMEQQREDPGKRLAHKRLAAEVTKLV  
HGKEGLDSAKRCTNVLYHNSVQALEEMSDGELQELFQEAPFHELLLEPGTTVIDACRQVD  
AIPAGHKGYLMVSEGAWINHTRVDKPEQVLI PKLHILSNGLTLIRVGKKNFSIIKWLSL  
480 aa / 18eme aa / 0,9885

>Yarovia.lypolitica-mito

MLRLRLVRNFTSKAGRISSTGTANVFRNEPKPANLQKSARDEAGFSSSDSAPSLQYLKD  
RNLVANVTEDSLEQQLQKFTSFYLGIDPSGSPMHLGHMVPVMIMLHLFLRGHYAFALVG  
GATGAVGDPTGKTEERKT VVKDTLQYNLDALATSLERTFLHAVAAAKKEGVYGDQIKLQ  
YAI VNNHDWWKNVGFLEFLASYGRLIRVNQMMARDSVKDRLNSATGIGYNEFSYQILQAF  
DFWHLFETQGCSLQLGGGDQWGNITAGVDLTKRAAAVNNLKETPYGITTTQLMLS KSGKKL  
GKSENNATIWLNPETRPFELYQYLLKTADEDVEQYLMFTFVSLEDIATIMRQHAENES  
LRYAQRVLADKFTALIHGDTVVGRCQVITRIMFAKSSLELNQQPKPAALTNSES LKQVLA  
EEGLLPLDKSTDYTTLVSTHFNMSSAKKLI SNKGVSVGLEGAQKVEIGPILEKHKDM  
GDFLIVRVGKLVQPFYLGSDPVETIPSAVGISDSYEVHTEDIDVTRK  
527 aa / 29eme aa / 0,9843

## Résumé en français

L'aminocyclation des ARN de transfert (ARNt) consiste en la fixation covalente d'un acide aminé sur l'extrémité 3' de l'ARNt correspondant par l'aminocycl-ARNt synthétase (aaRS) homologue. L'aminocycl-ARNt ainsi formé sera acheminé au ribosome pour traduire le codon de l'ARN messager correspondant à son anticodon et l'acide aminé sera inséré dans la chaîne polypeptidique en cours de synthèse. Chaque cellule est dotée d'une vingtaine d'aaRS, chacune devant être spécifique d'un acide aminé et d'une famille d'ARNt isoaccepteurs.

Le système d'aminocyclation spécifique de la tyrosine a été largement étudié ces 40 dernières années. La tyrosyl-ARNt synthétase (TyrRS) de *Bacillus stearothermophilus* est d'ailleurs la première aaRS dont la structure cristallographique a été résolue. Cependant, de nombreuses questions restent sans réponse aussi bien au niveau structural que fonctionnel et évolutif. De plus, des différences fonctionnelles significatives dans le mécanisme de reconnaissance entre l'ARNt<sup>Tyr</sup> et la TyrRS ont été décrites entre les bactéries et les archaëa/eucaryotes. En effet, contrairement à ce qui a été observé pour les autres systèmes d'aminocyclation, les réactions de charges croisées entre ces deux groupes sont impossibles. Ceci est dû d'une part à la nature différente de la première paire de bases 1-72 de l'ARNt<sup>Tyr</sup> et d'autre part à l'organisation particulière du site de fixation de la branche acceptrice de l'ARNt sur les TyrRS eubactériennes et archaëa/eucaryotiques.

Ma thèse porte sur l'étude fonctionnelle et structurale des partenaires de la réaction d'aminocyclation spécifique de la tyrosine dans la mitochondrie humaine. En effet, l'origine endosymbiotique et l'évolution de la mitochondrie en font un système hybride entre eubactérie et eucaryote. Au cours des deux premières années de ma thèse, j'ai caractérisé et analysé fonctionnellement la TyrRS mitochondriale humaine. Dans un premier temps, j'ai recherché informatiquement, cloné, surexprimé et purifié cette synthétase (**Art. 1 et 2**) ainsi que son ARNt homologue et certains variants. J'ai montré que la mutation de la première paire de bases (G1-C72) de l'ARNt<sup>Tyr</sup> n'a pas d'effet sur l'aminocyclation par la TyrRS mitochondriale. Ces résultats ont révélé une spécificité relâchée de l'enzyme, ce qui est également suggéré par l'analyse de la séquence primaire de l'enzyme. De plus, cette TyrRS est capable de fixer efficacement la tyrosine sur des ARNt<sup>Tyr</sup> d'organismes variés. De ce fait, la TyrRS mitochondriale humaine est la première TyrRS connue à ce jour qui s'affranchisse de la barrière d'espèces et qui ne discrimine pas les ARNt<sup>Tyr</sup> en fonction de la nature de leur première paire de bases (**Art. 3 et Rev. 1**).

Pour comprendre les particularités et les bases moléculaires de la réaction de tyrosylation dans la mitochondrie, l'étude structurale de la TyrRS a été réalisée au cours des deux années suivantes. J'ai cristallisé l'enzyme entière ainsi qu'une version minimale active (composée des domaines catalytique et de liaison à l'anticodon). Les cristaux de cette dernière forme ont diffracté les rayons X jusqu'à 2.2 Å et ont permis de construire une structure à haute résolution de l'enzyme en présence d'un analogue d'adénylate de tyrosine (**Art. 4 et 5**). La TyrRS mitochondriale est un homodimère de forme allongée avec les domaines de liaison à l'anticodon aux extrémités, et qui est susceptible de fixer l'ARNt<sup>Tyr</sup> à cheval sur ses deux monomères. L'analyse précise du site catalytique montre que le mécanisme de fixation de l'ATP et de la tyrosine est conservé chez toutes les TyrRS. Par ailleurs, l'enzyme mitochondriale se distingue des autres TyrRS par la présence de deux insertions à sa surface, l'une potentiellement impliquée dans la reconnaissance de l'ARNt<sup>Tyr</sup> et l'autre qui pourrait constituer une zone d'interaction avec un cofacteur. Enfin, la nature et la configuration du site de reconnaissance de la première paire de bases diffèrent par rapport à celles des autres TyrRS. Ces particularités sont à l'origine de la perte de spécificité de la TyrRS et sa capacité à charger les ARNt<sup>Tyr</sup> quelles que soient leurs origines. Ces différences sont probablement le reflet de l'origine endosymbiotique de la mitochondrie et de son contexte évolutif.

## Résumé en anglais

Aminoacylation of transfer RNA (tRNA), where a given amino acid is covalently bound to the 3' end of the cognate tRNA, is catalysed by the homologous aminoacyl-tRNA synthetase (aaRS). The generated aminoacyl-tRNA is driven to the ribosome for translation of the messenger RNA codon corresponding to its anticodon. This will result in the insertion of the amino acid in the synthesized polypeptide chain. Each cell possess about 20 aaRSs, each specific for an amino acid and a family of isoaccepting tRNAs.

The tyrosine-specific aminoacylation system has been deeply studied in the four past decades. The tyrosyl-tRNA synthetase (TyrRS) from *Bacillus stearothermophilus* is indeed the first aaRS which crystallographic structure has been solved. Still, many questions are unanswered at the functional, structural and evolutionary levels. Moreover, significant functional discrepancies between eubacteria and archaea/eukarya have been described for the recognition mechanism between tRNA<sup>Tyr</sup> and TyrRS. Indeed, unlike the general situation found in other aminoacylation systems, cross-reactions between those two groups are impossible. This is due to the difference of tRNA<sup>Tyr</sup> 1-72 first base pair and to the peculiar organisation of the tRNA acceptor end binding site of eubacterial and archaeal/eukaryal TyrRSs.

My thesis focuses on the functional and structural characterisation of the tyrosine specific aminoacylation system from human mitochondria. Mitochondrion is a hybrid between eubacteria and eukarya due to its endosymbiotic origin. During the first two years of my thesis, I characterised and functionally analysed the human mitochondrial TyrRS. First, I bioinformatically searched, cloned, overexpressed and purified this enzyme and its homologous tRNA (wild-type and mutated) (**Art. 1 & 2**). I showed that mutation of the tRNA<sup>Tyr</sup> first base pair has no effect on aminoacylation by the mitochondrial TyrRS. This relaxed specificity of the enzyme was also suggested by the analysis of its primary sequence. Further, this TyrRS is able to productively bind tRNA<sup>Tyr</sup> from various organisms (eubacteria, archaea and eukarya). Human mitochondrial is thus the first known TyrRS overcoming the species barrier and that doesn't discriminate tRNA<sup>Tyr</sup> for their first base pair (**Art. 3 & Rev. 1**).

To understand the peculiarities and the molecular bases of the tyrosylation reaction in human mitochondria, the structural study of this TyrRS was done during the two next years. I crystallised the full length enzyme and a minimal active version (consisting in the catalytic and anticodon binding domains). The crystals of this latter form diffracted X-rays at 2.2 Å resolution and I was able to build a structure of the enzyme in complex with a tyrosyl-adenylate analog at high resolution (**Art. 4 & 5**). Human mitochondrial TyrRS is an homodimer of elongated shape with its anticodon binding domains at both extremities and that will likely bind one tRNA<sup>Tyr</sup> molecule across both subunits. Precise analyse of the catalytic site reveals that the tyrosine and ATP binding mechanism is conserved among all TyrRSs. Further, the mitochondrial enzyme differs from other TyrRS by the presence of two insertions at its surface, the first one being potentially implicated in the tRNA<sup>Tyr</sup> binding and the second one that could serve as a co-factor binding interface. Finally, the nature and configuration of the tRNA<sup>Tyr</sup> first base pair recognition site differs from the ones from archaeal, eukaryal and eubacterial TyrRSs. These peculiarities are the roots for the loss of specificity and the ability to charge any tRNA<sup>Tyr</sup> in the case of human mitochondria. Such idiosyncrasies are probably due to the endosymbiotic origin of mitochondria and its evolutionary context.



## Références bibliographiques





# A

- Adams, P.D., Grosse-Kunstleve, R.W., Hung, L.W., Ioerger, T.R., McCoy, A.J., Moriarty, N.W., Read, R.J., Sacchettini, J.C., Sauter, N.K. et Terwilliger, T.C. (2002) PHENIX: building new software for automated crystallographic structure determination. *Acta Cryst.*, **D58**, 1948-1954.
- Ahel, I., Korencic, D., Ibba, M. et Söll, D. (2003) Trans-editing of mischarged tRNAs. *Proc. Natl. Acad. Sci. U.S.A.*, **100**, 15422-15427.
- Akins, R.A. et Lambowitz, A.M. (1987) A protein required for splicing group I introns in *Neurospora* mitochondria is mitochondrial tyrosyl-tRNA synthetase or a derivative thereof. *Cell*, **50**, 331-345.
- Allmang, C. et Krol, A. (2006) Selenoprotein synthesis: UGA does not end the story. *Biochimie*, **88**, 1561-1571.
- Ambrogelly, A., Korencic, D. et Ibba, M. (2002) Functional annotation of class I lysyl-tRNA synthetase phylogeny indicates a limited role for gene transfer. *J. Bacteriol.*, **184**, 4594-4600.
- Ames, B.N., Tsang, T.H., Buck, M. et Christman, M.F. (1983) The leader mRNA of the histidine attenuator region resembles tRNA<sup>His</sup>; possible general regulatory implications. *Proc. Natl. Acad. Sci. USA*, **80**, 5240-5242.
- An, S. et Musier-Forsyth, K. (2004) Trans-editing of Cys-tRNA<sup>Pro</sup> by Haemophilus influenzae YbaK protein. *J. Biol. Chem.*, **279**, 42359-42362.
- Anderson, S., Bankier, A.T., Barrell, B.G., de Bruijn, M.H.L., Coulson, A.R., Drouin, J., Eperon, J.C., Nierlich, D.P., Roe, B.A., Sanger, F., Schreier, P.H., Smith, A.J.H., Staden, R. et Young, I.G. (1981) Sequence and organization of the human mitochondrial genome. *Nature*, **290**, 457-465.
- Antonellis, A., Ellsworth, R.E., Sambuughin, N., Puls, I., Abel, A., Lee-Lin, S.Q., Jordanova, A., Kremensky, I., Christodoulou, K., Middleton, L.T., Sivakumar, K., Ionasescu, V., Funalot, B., Vance, J.M., Goldfarb, L.G., Fischbeck, K.H. et Green, E.D. (2003) Glycyl tRNA synthetase mutations in Charcot-Marie-Tooth disease type 2D and distal spinal muscular atrophy type V. *Am. J. Hum. Genet.*, **72**, 1293-1299.
- Antonellis, A., Lee-Lin, S.-Q., Wasterlain, A., Leo, P., Quezado, M., Goldfarb, L.G., Myung, K., Burgess, S., Fischbeck, K.H. et Green, E.D. (2006) Functional analyses of glycyl-tRNA synthetase mutations suggest a key role for tRNA-charging enzymes in peripheral axons. *J. Neurosci.*, **26**, 10397-10406.
- Arnez, J. et Cavarelli, J. (1997) Structure of RNA-binding proteins. *Q. Rev. Biophys.*, **30**, 195-240.
- Attardi, G. et Montoya, J. (1983) Analysis of human mitochondrial RNA. *Methods Enzymol.*, **97**, 435-469.

Augustin, S., Nolden, M., Müller, S., Hardt, O., Arnold, I. et Langer, T. (2005) Characterization of peptides released from mitochondria: evidence for constant proteolysis and peptide efflux. *J. Biol. Chem.*, **280**, 2691-2699.

## B

Baker, N.A., Sept, D., Joseph, S., Holst, M.J. et McCammon, J.A. (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U.S.A.*, **98**, 10037-10041.

Baldwin, A.N. et Berg, P. (1966) Transfer ribonucleic acid-induced hydrolysis of valyladenylate bound to isoleucyl ribonucleic acid synthetase. *J. Biol. Chem.*, **241**, 839-845.

Barends, S., Bink, H.H., Van den Worm, S.H.E., Pleij, C.W. et Kraal, B. (2003) Entrapping ribosomes for viral translation: tRNA mimicry as a molecular trojan horse. *Cell*, **112**, 123-129.

Baron, C., Westhof, E., Böck, A. et Giegé, R. (1993) Solution structure of selenocysteine inserting tRNA<sup>Sec</sup> from *Escherichia coli*. Comparison with canonical tRNA<sup>Ser</sup>. *J. Mol. Biol.*, **231**, 274-292.

Becker, H.D., Giegé, R. et Kern, D. (1996) Identity of prokaryotic and eukaryotic tRNA<sup>Asp</sup> for aminoacylation by aspartyl-tRNA synthetase from *Thermus thermophilus*. *Biochemistry*, **35**, 7447-7458.

Becker, H.D. et Kern, D. (1998) *Thermus thermophilus* - a link in evolution of the tRNA-dependent amino acid amidation pathways. *Proc. Natl. Acad. Sci.*, **95**, 12832-12837.

Bedouelle, H. et Winter, G. (1986) A model of synthetase/transfer RNA interaction as deduced by protein engineering. *Nature*, **320**, 371-373.

Beebe, K., Merriman, E., Pouplana, L.R.D. et Schimmel, P. (2004) A domain for editing by an archaeobacterial tRNA synthetase. *Proc. Natl. Acad. Sci. U.S.A.*, **101**, 5958-5963.

Bernier, S., Dubois, D.Y., Therrien, M., Lapointe, J. et Chenevert, R. (2000) Synthesis of glutaminyl adenylate analogues that are inhibitors of glutaminyl-tRNA synthetase. *Bioorg. Med. Chem. Lett.*, **10**, 2441-2444.

Betteridge, Z., Gunawardena, H., North, J., Slinn, J. et McHugh, N. (2007) Anti-synthetase syndrome: a new autoantibody to phenylalanyl transfer RNA synthetase (anti-Zo) associated with polymyositis and interstitial pneumonia. *Rheumatology (Oxford)*.

Beuning, P.J. et Musier-Forsyth, K. (1999) Transfer RNA recognition by aminoacyl-tRNA synthetases. *Biopolymers*, **52**, 1-28.

Bhattacharyya, S.N. et Adhya, S. (2004) The complexity of mitochondrial tRNA import. *RNA Biol.*, **1**, 84-88.

Blobstein, S.H., Grunberger, D., Weinstein, I.B. et Nakanishi, K. (1973) Isolation and structure determination of the fluorescent base from bovine liver phenylalanine transfer ribonucleic acid. *Biochemistry*, **12**, 188-193.

- Brevet, A., Chen, J., Leveque, F., Blanquet, S. et Plateau, P. (1995) Comparison of the enzymatic properties of the two *Escherichia coli* lysyl-tRNA synthetase species. *J. Biol. Chem.*, **270**, 14439-14444.
- Brick, P., Bhat, T.N. et Blow, D.M. (1989) Structure of tyrosyl-tRNA synthetase refined at 2.3 Å resolution. Interaction of the enzyme with tyrosyl adenylate intermediate. *J. Mol. Biol.*, **208**, 83-98.
- Brick, P. et Blow, D.M. (1987) Crystal structure of a deletion mutant of a tyrosyl-tRNA synthetase complexed with tyrosine. *J. Mol. Biol.*, **194**, 287-297.
- Brindefalk, B., Viklund, J., Larsson, D., Thollesson, M. et Andersson, S.G. (2007) Origin and Evolution of the Mitochondrial Aminoacyl-tRNA Synthetases. *Mol. Biol. Evol.*, **24**, 743-756.
- Brulé, H., Holmes, W.M., Keith, G., Giegé, R. et Florentz, C. (1998) Effect of a mutation in the anticodon of human mitochondrial tRNA<sup>Pro</sup> on its post-transcriptional modification pattern. *Nucleic Acids Res.*, **26**, 537-543.
- Brunel, C. et Romby, P. (2000) Probing RNA structure and RNA-ligand complexes with chemical probes. *Methods Enzymol.*, **318**, 3-21.
- Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T. et Warren, G.L. (1998) Crystallography and NMR system: A new software suite for macromolecular structure determination. *Acta Cryst.*, **D54**, 905-921.
- Buck, C.A. et Nass, M.M.K. (1969) Studies on mitochondrial tRNA from animal cells. I. A comparison of mitochondrial and cytoplasmic tRNA and aminoacyl-tRNA synthetases. *J. Mol. Biol.*, **41**, 67-82.
- Buddha, M.R., Keery, K.M. et Crane, B.R. (2004) An unusual tryptophanyl tRNA synthetase interacts with nitric oxide synthase in *Deinococcus radiodurans*. *Proc. Natl. Acad. Sci.*, **101**, 15881-15886.
- Bullard, J.-M., Cai, Y.-C. et Spremulli, L.-L. (2000) Expression and characterization of the human mitochondrial leucyl-tRNA synthetase. *Biochem. Biophys. Acta*, **1490**, 245-258.
- Bullard, J., Cai, Y.-C., Demeler, B. et Spremulli, L. (1999) Expression and characterization of a human mitochondrial phenylalanyl-tRNA synthetase. *J. Mol. Biol.*, **288**, 567-577.
- Bunn, C.C., Bernstein, R.M. et Mathews, M.B. (1986) Autoantibodies against alanyl-tRNA synthetase and tRNA<sup>Ala</sup> coexist and are associated with myositis. *J. Exp. Med.*, **163**, 1281-1291.

## C

- Calendar, R. et Berg, P. (1967) D-Tyrosyl RNA: formation, hydrolysis and utilization for protein synthesis. *J. Mol. Biol.*, **26**, 39-54.

- Caprara, M.G., Lehnert, V., Lambowitz, A.M. et Westhof, E. (1996) A tyrosyl-tRNA synthetase recognizes a conserved tRNA-like structural motif in the group I intron catalytic core. *Cell*, **87**, 1135-1145.
- Cathopoulis, T., Chuawong P. et Hendrickson T.L. (2007) Novel tRNA aminoacylation mechanisms. *Mol. Biosyst.*, **3**, 408-418.
- Cen, S., Javanbakht, H., Niu, M. et Kleiman, L. (2004) Ability of wild-type and mutant lysyl-tRNA synthetase to facilitate tRNA<sup>Lys</sup> incorporation into human immunodeficiency virus type 1. *J. Virol.*, **78**, 1595-1601.
- Cen, S., Khorchid, A., Javanbakht, H., Gabor, J., Stello, T., Shiba, K., Musier-Forsyth, K. et Kleiman, L. (2001) Incorporation of Lysyl-tRNA synthetase into human immunodeficiency Virus Type 1. *J. Virol.*, **75**, 5043-5048.
- Chimnarong, S., Jeppesen, M.G., Suzuki, T., Nyborg, J. et Watanabe, K. (2005) Dual-mode recognition of noncanonical tRNA<sup>Ser</sup> by seryl-tRNA synthetase in mammalian mitochondria. *EMBO J.*, **24**, 3369-3379.
- Claros, M.G. et Vincens, P. (1996) Computational method to predict mitochondrially imported proteins and their targeting sequences. *Eur. J. Biochem.*, **241**, 779-786.
- Claverie, J.-M. et Notredame, C. (2006) *Bioinformatics For Dummies, 2nd Edition*
- Clayton, D.A. (1982) Replication of animal mitochondrial DNA. *Cell*, **28**, 693-705.
- Collaborative Computational Project, N., 1994. (1994) The *CCP4* Suite: programs for protein crystallography. *Acta Cryst.*, **D50**, 760-763.
- Crick, F.H.C. (1955) On degenerate templates and the adaptor hypothesis: A note for the RNA tie club; non publié mais cité par Haogland M.B. (1960); dans the *Nucleic Acids* (eds) Chargaff E. et Davidson J.N. (New York: Academic Press) Vol 3, p 349
- Cusack, S., Berthet-Colominas, C., Härtlein, M., Nassar, N. et Leberman, R. (1990) A second class of synthetase structure revealed by X-ray analysis of *Escherichia coli* seryl-tRNA synthetase at 2.5 Å. *Nature*, **347**, 249-255.

## D

- Dagkessamanskaia, A., Martin-Yken, H., Basmaji, F., Briza, P. et Francois, J. (2001) Interaction of Knr4 protein, a protein involved in cell wall synthesis, with tyrosine tRNA synthetase encoded by *TYS1* in *Saccharomyces cerevisiae*. *FEMS Microbiol. Lett.*, **200**, 53-58.
- Dardel, F., Panvert, M. et Fayat, G. (1990) Transcription and regulation of expression of the *Escherichia coli* methionyl-tRNA synthetase gene. *Mol. Gen. Genet.*, **223**, 121-133.
- de Duve, C. (1988) The second genetic code. *Nature*, **333**, 117-118.
- Degoul, F., Brulé, H., Capanec, C., Helm, M., Marsac, C., Leroux, J.-P., Giegé, R. et Florentz, C. (1998) Isoleucylation properties of native human mitochondrial tRNA<sup>Ile</sup> and tRNA<sup>Ile</sup> transcripts. Implications for cardiomyopathy-related point mutations (4269, 4317) in the tRNA<sup>Ile</sup> gene. *Hum. Mol. Gen.*, **7**, 347-354.

- DeLano, W.L. (2002) The PyMOL molecular graphics system. Delano Scientific (<http://www.pymol.org>), San Carlos, CA, USA.
- Delarue, M., Poterszman, A., Nikonov, S., Garber, M., Moras, D. et Thierry, J.C. (1994) Crystal structure of a prokaryotic aspartyl tRNA-synthetase. *EMBO J.*, **13**, 3219-3229.
- Deutscher, M.P. (1983) tRNA nucleotidyl-transferase and the CCA terminus of transfer RNA. In Jacob, S.T. (ed.), *Enzymes of nucleic acid synthesis and modification*. CRC Press, Inc., Boca Raton, FL, USA, Vol. 2, pp. 159-183.
- Di Giulio, M. (1992) On the origin of the transfer RNA molecule. *J. Theor. Biol.*, **159**, 199-214.
- Di Giulio, M. (1999) The non-monophyletic origin of the tRNA molecule. *J. Theor. Biol.*, **197**, 403-414.
- Di Giulio, M. (2004) The origin of the tRNA molecule: implications for the origin of protein synthesis. *J. Theor. Biol.*, **226**, 89-93.
- Diaz-Lazcoz, Y., Aude, J.C., Nitschké, P., Chiapello, H., Landès-Devauchelle, C. et Risler, J.L. (1998) Evolution of genes, evolution of species: the case of aminoacyl-tRNA synthetases. *Mol. Biol. Evol.*, **15**, 1548-1561.
- Dickey, T., Khaki, A., Gauger, A., Nieuwkoop, A., Agyiri, A., Ducker, G. et Chihade, J. (2006) Animal Mitochondrial Alanine-tRNA Synthetase Use Divergent tRNA Recognition Strategies. *2006 International Conference on Aminoacyl-tRNA Synthetases: From the Genetic Code to Human Diseases & Medicine*.
- DiMauro, S. et Moraes, C.T. (1993) Mitochondrial encephalomyopathies. *Arch. Neurol.*, **50**, 1197-1208.
- DiMauro, S. et Schon, E. (2001) Mitochondrial DNA Mutations in Human Disease. *Am. J. Med. Genet.*, **106**, 18-26.
- Dock-Bregeon, A.-C., Rees, B., Torres-Larios, A., Bey, G., Caillet, J. et Moras, D. (2004) Achieving error-free translation; the mechanism of proofreading of threonyl-tRNA synthetase at atomic resolution. *Mol. Cell*, **16**, 375-386.
- Dock-Bregeon, A., Sankaranarayanan, R., Romby, P., Caillet, J., Springer, M., Rees, B., Francklyn, C.S., Ehresmann, C. et Moras, D. (2000) Transfer RNA-mediated editing in threonyl-tRNA synthetase. The class II solution to the double discrimination problem. *Cell*, **103**, 877-884.
- Dörner, M., Altmann, M., Pääbo, S. et Mörl, M. (2001) Evidence for import of a lysyl-tRNA into marsupial mitochondria. *Mol. Biol. Cell.*, **12**, 2688-2698.
- Dreher, T.W. (1999) Functions of 3'-untranslated regions of positive strand RNA viral genomes. *Annu. Rev. Phytopathol.*, **37**, 151-174.
- Dubois, D.Y., Blaise, M., Becker, H.D., Campanacci, V., Keith, G., Giegé, R., Cambillau, C., Lapointe, J. et Kern, D. (2004) An aminoacyl-tRNA synthetase-like protein encoded by the *Escherichia coli* yadB gene glutamylates specifically tRNA<sup>Asp</sup>. *Proc. Natl. Acad. Sci. U.S.A.*, **101**, 7530-7535.

Ducruix, A. et Giegé, R. (1999) Methods of crystallization. In Ducruix, A. et Giegé, R. (eds.), *Crystallization of Nucleic Acids and Proteins: A Practical Approach*. IRL Press, Oxford, pp. 121-147.

Dunn, J.J. et Studier, F.W. (1983) Complete nucleotide sequence of bacteriophage T7 DNA and the location of T7 genetic elements. *J. Mol. Biol.*, **166**, 477-535.

## E

Ebel, J.-P., Giegé, R., Bonnet, J., Kern, D., Befort, N., Bollack, C., Fasiolo, F., Gangloff, J. et Dirheimer, G. (1973) Factors determining the specificity of the tRNA aminoacylation reaction. *Biochimie*, **55**, 547-557.

Eldred, E.W. et Schimmel, P. (1972) Investigation of the transfer of amino acid from a tRNA synthetase-aminoacyl adenylate complex to tRNA. *Biochemistry*, **11**, 17-23.

Ellman, G.L. (1959) Tissue sulfhydryl groups. *Arch. Biochem. Biophys.*, **82**, 70-77.

Emsley, P. et Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Cryst.*, **D60**, 2126-2132.

Eriani, G., Delarue, M., Poch, O., Gangloff, J. et Moras, D. (1990) Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs. *Nature*, **347**, 203-206.

## F

Faulhammer, H.G. et Cramer, F. (1977) Tyrosyl-tRNA synthetase from baker's yeast. Rapid isolation by affinity elution, molecular weight of the enzyme, and determination of essential sulfhydryl groups. *Eur. J. Biochem.*, **75**, 561-570.

Fechter, P., Rudinger-Thirion, J., Florentz, C. et Giegé, R. (2001) Novel features in the tRNA-like world of plant viral RNAs. *Cell. Mol. Life Sci.*, **58**, 1547-1561.

Fechter, P., Rudinger, J., Giegé, R. et Théobald-Dietrich, A. (1998) Ribozyme processed tRNA transcripts with unfriendly internal promoter for T7 RNA polymerase: production and activity. *FEBS Lett.*, **436**, 99-103.

Fender, A., Sauter, C., Messmer, M., Pütz, J., Giegé, R., Florentz, C. et Sissler, M. (2006) Loss of a primordial identity element for a mammalian mitochondrial aminoacylation system. *J. Biol. Chem.*, **281**, 15980-15986.

Ferreira, R. et Cavalcanti, A.R. (1997) Vestiges of early molecular processes leading to the genetic code. *Orig. Life Evol. Biosph.*, **27**, 397-403.

Ferri-Fioni, M.-L., Fromant, M., Bouin, A.-P., Aubard, C., Lazennec, C., Plateau, P. et Blanquet, S. (2006) Identification in archaea of a novel D-Tyr-tRNA<sup>Tyr</sup> deacylase. *J. Biol. Chem.*, **281**, 27575-27585.

Fersht, A. (1977) *Enzyme, Structure and Mechanism*. W. H. Freeman and Company, Reading and San Francisco.

- Fersht, A.R. (1987) The hydrogen bond in molecular recognition. *Trends in Biochem. Sci.*, **12**, 301-304.
- Florentz, C., Sohm, B., Tryoen-Tóth, P., Pütz, J. et Sissler, M. (2003) Human mitochondrial tRNAs in health and disease. *Cell. Mol. Life Sci.*, **60**, 1356-1375.
- Förster, C., Limmer, S., Zeidler, W. et Sprinzl, M. (1994) Effector region of the translation elongation factor EF-Tu-GTP complex stabilizes an orthoester acid intermediate structure of aminoacyl-tRNA in a ternary complex. *Proc. Natl. Acad. Sci. U.S.A.*, **92**, 4254-4257.
- Francklyn, C. et Schimmel, P. (1989) Aminoacylation of RNA minihelices with alanine. *Nature*, **337**, 478-481.
- Frugier, M., Florentz, C. et Giegé, R. (1994) Efficient aminoacylation of resected RNA helices by class II aspartyl-tRNA synthetase dependent on a single nucleotide. *EMBO J.*, **13**, 2218-2226.
- Frugier, M. et Giegé, R. (2003) Yeast aspartyl-tRNA synthetase binds specifically its own mRNA. *J. Mol. Biol.*, **331**, 375-383.
- Frugier, M., Helm, M., Felden, B., Giegé, R. et Florentz, C. (1998) Sequences outside recognition sets are not neutral for tRNA aminoacylation: evidence for non-permissive combinations of nucleotides in the acceptor stem of yeast tRNA<sup>Phe</sup>. *J. Biol. Chem.*, **273**, 11605-11610.
- Frugier, M., Ryckelynck, M. et Giegé, R. (2005) tRNA-balanced expression of a eukaryal aminoacyl-tRNA synthetase by an mRNA-mediated pathway. *EMBO Rep.*, **6**, 860-865.
- Fukai, S., Nureki, O., Sekine, S., Shimada, A., Tao, J., Vassylyev, D.G. et Yokoyama, S. (2000) Structural basis for double-sieve discrimination of L-valine from L-isoleucine and L-threonine by the complex of tRNA<sup>Val</sup> and valyl-tRNA synthetase. *Cell*, **103**, 793-803.
- Fukunaga, R. et Yokoyama, S. (2007a) Structure of the AlaX-M trans-editing enzyme from *Pyrococcus horikoshii*. *Acta Cryst.*, **D63**, 390-400.

## G

- Gabriel, K. et McClain, W.H. (1999) A set of plasmids constitutively producing different RNA levels in *Escherichia coli*. *J. Mol. Biol.*, **290**, 385-390.
- Galtier, N., Gouy, M. et Gautier, C. (1996) SEAVIEW and PHYLO\_WIN: two graphic tools for sequence alignment and molecular phylogeny. *Comput. Appl. Biosci.*, **12**, 543-548.
- Giegé, R. et Frugier, M. (2003) Transfer RNA structure and identity. In Lapointe, J. et Brakier-Gringas, L. (eds.), *Translation mechanisms*. Landes Sciences, Georgetown, TX, pp. 1-24.
- Giegé, R., Helm, M. et Florentz, C. (2001) Classical and novel chemical tools for RNA structure probing. In Söll, D., Nishimura, S. et Moore, P. (eds.), *RNA*. Elsevier Science B.V., Amsterdam, pp. 71-89.

- Giegé, R., Sissler, M. et Florentz, C. (1998) Universal rules and idiosyncratic features in tRNA identity. *Nucleic Acids Res.*, **26**, 5017-5035.
- Gilbert, W., de Souza, S.J. et Long, M. (1997) Origin of genes. *Proc. Natl. Acad. Sci. U.S.A.* , **94**, 7698-7703.
- Glaser, P., Kunst, F., Debarbouille, M., Vertes, A., Danchin, A. et Dedonder, R. (1991) A gene encoding a tyrosine tRNA synthetase is located near *sacS* in *Bacillus subtilis*. *DNA Seq.*, **1**, 251-261.
- Godinic, V., Mocibob, M., Rocak, S., Ibba, M. et Weygand-Durasevic, I. (2007) Peroxin Pex21p interacts with the C-terminal noncatalytic domain of yeast seryl-tRNA synthetase and forms a specific ternary complex with tRNA<sup>Ser</sup>. *FEBS J.*, **in press**.
- Gouet, P. et Courcelle, E. (2002) ENDscript: a workflow to display sequence and structure information. *Bioinformatics*, **18**, 767-768.
- Gouet, P., Courcelle, E., Stuart, D.I. et Metz, F. (1999) ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics*, **15**, 305-308.
- Gouet, P., Robert, X. et Courcelle, E. (2003) ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res.*, **31**, 3320-3323.
- Graffe, M., Dondon, J., Caillet, J., Romby, P., Ehresmann, C., Ehresmann, B. et Springer, M. (1992) The specificity of translational control switched using tRNA identity rules. *Science*, **255**, 994-996.
- Gregson, J.M., Crain, P.F., Edmonds, C.G., Gupta, R., Hashizume, T., Philipson, D.W. et McCloskey, J.A. (1993) Structure of the archaeal transfer RNA nucleoside G<sup>-15</sup> (2-amino-4,7-dihydro-4-oxo-7-beta-D-ribofuranosyl-1H-pyrrolo[2,3-d]pyrimidine-5-carboximide (archaeosine)). *J. Biol. Chem.*, **268**, 10076-10086.
- Grosjean, H. et Benne, R. (1998) *Modification and Editing of RNA*. ASM Press, Washington.
- Guijarro, J.I., Pintar, A., Prochnicka-Chalufour, A., Guez, V., Gilquin, B., Bedouelle, H. et Delepierre, M. (2002) Structure and dynamics of the anticodon arm binding domain of *Bacillus stearothermophilus* tyrosyl-tRNA synthetase. *Structure*, **10**, 311-317.
- Guindon, S. et Gascuel, O. (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.*, **52**, 696-704.

## H

- Hammen, P.K. et Weiner, H. (1998) Mitochondrial leader sequences: structural similarities and sequence differences. *J. Exp. Zool.* , **282**, 280-283.
- Han, J.M., Lee, M.J., Park, S.G., Lee, S.H., Razin, E., Choi, E.-C. et Kim, S. (2006) Hierarchical Network between the Components of the Multi-tRNA Synthetase Complex: Implications for Complex Formation. *J. Biol. Chem.*, **281**, 38663--38667.



- Hao, H. et Moraes, C.T. (1997) A disease-associated G5703A mutation in human mitochondrial DNA causes a conformational change and a marked decrease in steady-state levels of mitochondrial tRNA<sup>Asn</sup>. *Mol. Cel. Biol.*, **17**, 6831-6837.
- Hashish, L., Trieu, E.P., Sadanandan, P. et Targoff, I.N. (2005) Identification of autoantibodies to tyrosyl-tRNA synthetase in dermatomyositis with features consistent with antisynthetase syndrome. *Arthritis Rheum.*, **52**, S312.
- Hayashi, I., Kawai, G. et Watanabe, K. (1998) Higher-order structure and thermal instability of bovine mitochondrial tRNA<sup>Ser</sup><sub>UGA</sub> investigated by proton NMR spectroscopy. *J. Mol. Biol.*, **284**, 57-69.
- Heckman, J.E., Alzner-Deweerd, B. et RajBhandary, U.L. (1979) Interesting and unusual features in the sequence of *Neurospora crassa* mitochondrial tyrosine transfer RNA. *Proc. Natl. Acad. Sci. U.S.A.*, **76**, 717-721.
- Helm, M., Brulé, H., Friede, D., Giegé, R., Pütz, J. et Florentz, C. (2000) Search for characteristic structural features of mammalian mitochondrial tRNAs. *RNA*, **6**, 1356-1379.
- Helm, M., Giegé, R. et Florentz, C. (1999) A Watson-Crick base-pair disrupting methyl group (m<sup>1</sup>A9) is sufficient for cloverleaf folding of human mitochondrial tRNA<sup>Lys</sup>. *Biochemistry*, **38**, 13338-13346.
- Hendrickson, T.L., Nomanbhoy, T.K., de Crécy-Lagard, V., Fukai, S., Nureki, O., Yokoyama, S. et Schimmel, P. (2002) Mutational separation of two pathways for editing by a class I tRNA synthetase. *Mol. Cell*, **9**, 353--362.
- Himeno, H., Asahara, H., Tamura, K., Hasegawa, T., Ueda, T., Watanabe, K. et Shimizu, M. (1990) Discriminator base and anticodon as a tRNA identity element. *Nucleic Acids Symp. Ser.*
- Hirakata, M., Satoh, S. et Mimori, T. (1999) [Anti-PL-7 autoantibodies and anti-PL-12 autoantibodies (anti-aminoacyl-tRNA 'other than histidyl' synthetase autoantibodies)]. *Nippon Rinsho*, **57 Suppl**, 435-438.
- Hoagland, M.B. (1955) An enzymic mechanism for amino acid activation in animal tissues. *Biochem. Biophys. Acta*, **16**, 288-289.
- Hoagland, M.B., Stephenson, M.L., Scott, J.F., Hecht, L.I. et Zamecnik, P.C. (1958) A soluble ribonucleic acid intermediate in protein synthesis. *J. Biol. Chem.*, **231**, 241-257.
- Hoagland, M.B., Zamecnik, P.C. et Stephenson, M.L. (1957) Intermediate reactions in protein biosynthesis. *Biochim. Biophys. Acta*, **24**, 215-216.
- Holley, R.W., Apgar, J., Everett, G.A., Madison, J.T., Marquise, M., Merrill, S.H., Penswick, J.R. et Zamir, R. (1965) Structure of a ribonucleic acid. *Science*, **147**, 1462-1465.
- Holt, I.J., Lorimer, H.E. et Jacobs, H.T. (2000) Coupled leading- and lagging-strand synthesis of mammalian mitochondrial DNA. *Cell*, **100**, 515-524.
- Horie, N., Hara-Yokoyama, M., Yokoyama, S., Watanabe, K., Kuchino, Y., Nishimura, S. et Miyazawa, T. (1985) Two tRNA<sup>Ile1</sup> species from an extreme thermophile, *Thermus*

*thermophilus* HB8: Effect of 2-thiolation of ribothymidine on the thermostability of tRNA. *Biochemistry*, **24**, 5711-5715.

Hou, Y.-M. et Schimmel, P. (1988) A simple structural feature is a major determinant of the identity of a transfer RNA. *Nature*, **333**, 140-145.

Hurdle, J.G., O'Neill, A.J. et Chopra, I. (2005) Prospects for Aminoacyl-tRNA Synthetase Inhibitors as New Antimicrobial Agents. *Antimicrob. Agents Chemother.*, **49**, 4821-4833.

Hussain, T., Kruparani, S.P., Pal, B., Dock-Bregeon, A.-C., Dwivedi, S., Shekar, M.R., Sureshbabu, K. et Sankaranarayanan, R. (2006) Post-transfer editing mechanism of a D-aminoacyl-tRNA deacylase-like domain in threonyl-tRNA synthetase from archaea. *EMBO J.*

## I

Ibba, M., Bono, J.L., Rosa, P.A. et Söll, D. (1997) Archaeal-type lysyl-tRNA synthetase in the Lyme disease spirochete *Borrelia burgdorferi*. *Proc. Natl. Acad. Sci. U.S.A.*, **94**, 14383-14388.

Ingman, M., Kaessmann, H., Pääbo, S. et Gyllenstein, U. (2000) Mitochondrial genome variation and the origin of modern humans. *Nature*, **408**, 708-713.

Ishitani, R., Nureki, O., Nameki, N., Okada, N., Nishimura, S. et Yokoyama, S. (2003) Alternative tertiary structure of tRNA for recognition by a posttranscriptional modification enzyme. *Cell*, **113**, 383-394.

Ivens, A., Mayans, O., Szadkowski, H., Jorgens, C., Wilmanns, M. et Kirschner, K. (2002) Stabilization of a (betaalpha)<sub>8</sub>-barrel protein by an engineered disulfide bridge. *Eur. J. Biochem.*, **269**, 1145-1153.

## J

James, B.D., Olsen, G.J., Liu, J. et Pace, N.R. (1988) The secondary structure of ribonuclease P RNA, the catalytic element of a ribonucleoprotein enzyme. *Cell*, **52**, 19-26.

Jordanova, A., Irobi, J., Thomas, F.P., Dijck, P.V., Meerschaert, K., Dewil, M., Dierick, I., Jacobs, A., Vriendt, E.D., Guerguelcheva, V., Rao, C.V., Tournev, I., Gondim, F.A.A., D'Hooghe, M., Gerwen, V.V., Callaerts, P., Bosch, L.V.D., Timmermans, J.-P., Robberecht, W., Gettemans, J., Thevelein, J.M., Jonghe, P.D., Kremensky, I. et Timmerman, V. (2006) Disrupted function and axonal distribution of mutant tyrosyl-tRNA synthetase in dominant intermediate Charcot-Marie-Tooth neuropathy. *Nat Genet.*, **38**, 197-202.

Jørgensen, R., Søgarrd, M.M., Rossing, A.B., Martensen, P.M. et Justesen, J. (2000) Identification and characterization of human mitochondrial tryptophanyl-tRNA synthetase. *J. Biol. Chem.*, **275**, 16820-16826.

Jovine, L., Djordjevic, S. et Rhodes, D. (2000) The crystal structure of yeast phenylalanine tRNA at 2.0 Å resolution: cleavage by Mg<sup>2+</sup> in 15-year old crystals. *J. Mol. Biol.*, **301**, 401-414.

## K

Kamiryo, T. et Matsubashi, M. (1969) Sequential addition of glycine from glycyl-tRNA to the lipid-linked precursors of cell wall peptidoglycan in *Staphylococcus aureus*. *Biochem. Biophys. Res. Commun.*, **36**, 215-222.

Kamper, U., Kück, U., Cherniack, A.D. et Lambowitz, A.M. (1992) The mitochondrial tyrosyl-tRNA synthetase of *Podospira anserina* is a bifunctional enzyme active in protein synthesis and RNA splicing. *Mol. Cell Biol.*, **12**, 499-511.

Kamtekar, S., Hohn, M.J., Park, H.-S., Schnitzbauer, M., Sauerwald, A., Söll, D. et Steitz, T.A. (2007) Toward understanding phosphoseryl-tRNA<sup>Cys</sup> formation: The crystal structure of *Methanococcus maripaludis* phosphoseryl-tRNA synthetase. *Proc. Natl. Acad. Sci. U.S.A.*, **104**, 2620-2625.

Keane, T.M., Creevey, C.J., Pentony, M.M., Naughton, T.J. et McLnerney, J.O. (2006) Assessment of methods for amino acid matrix selection and their use on empirical data shows that ad hoc assumptions for choice of matrix are not justified. *BMC Evol. Biol.*, **6**, 29.

Keiler, K.C., Waller, R.H. et Sauer, R.T. (1996) Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. *Science*, **271**, 990-993.

Kellermann, O., Tonetti, H., Brevet, A., Mirande, M., Pailliez, J.P. et Waller, J.P. (1982) Macromolecular complexes from sheep and rabbit containing seven aminoacyl-tRNA synthetases. I. Species specificity of the polypeptide composition. *J. Biol. Chem.*, **257**, 11041-11048.

Kelley, S., Steinberg, S. et Schimmel, P. (2000) Functional defects of pathogenic human mitochondrial tRNAs related to structural fragility. *Nature Struct. Biol.*, **7**, 862-865.

Kim, J.Y., Kang, Y.S., Lee, J.W., Kim, H.J., Ahn, Y.H., Park, H., Ko, Y.G. et Kim, S. (2002) p38 is essential for the assembly and stability of macromolecular tRNA synthetase complex: implications for its physiological significance. *Proc. Natl. Acad. Sci. U.S.A.*, **99**, 7912-7916.

Kim, S., Lee, S.W., Choi, E.C. et Choi, S.Y. (2003) Aminoacyl-tRNA synthetases and their inhibitors as a novel family of antibiotics. *Appl. Microbiol. Biotechnol.*, **61**, 278-288.

Kim, S.H., Suddath, F.L., Quigley, G.J., McPherson, A., Sussman, J.L., Wang, A.H.J., Seeman, N.C. et Rich, A. (1974) Three dimensional tertiary structure of yeast phenylalanine transfer RNA. *Science*, **185**, 435-440.

Kitabatake, M., Ali, K., Demain, A., Sakamoto, K., Yokoyama, S. et Soll, D. (2002) Indolmycin resistance of *Streptomyces coelicolor* A3(2) by induced expression of one of its two tryptophanyl-tRNA synthetases. *J. Biol. Chem.*, **277**, 23882-23887.

- Klipcan, L. et Safro, M. (2004) Amino acid biogenesis, evolution of the genetic code and aminoacyl-tRNA synthetases. *J. Theor. Biol.*, **228**, 389-396.
- Knapp, G., Ogden, R.C., Peebles, C.L. et Abelson, J. (1979) Splicing of yeast tRNA precursors: structure of the reaction intermediates. *Cell*, **18**, 37-45.
- Ko, Y.G., Kim, E.Y., Kim, T., Park, H., Park, H.S., Choi, E.J. et Kim, S. (2001) Glutamine-dependent antiapoptotic interaction of human glutaminyl-tRNA synthetase with apoptosis signal-regulating kinase 1. *J. Biol. Chem.*, **276**, 6030-6036.
- Kobayashi, T., Nureki, O., Ishitani, R., Yaremchuk, A., Tukalo, M., Cusack, S., Sakamoto, K. et Yokoyama, S. (2003) Structural basis for orthogonal tRNA specificities of tyrosyl-tRNA synthetases for genetic code expansion. *Nat. Struct. Biol.*, **10**, 425-432.
- Kobayashi, T., Takimura, T., Sekine, R., Vincent, K., Kamata, K., Sakamoto, K., Nishimura, S. et Yokoyama, S. (2005a) Structural Snapshots of the KMSKS Loop Rearrangement for Amino Acid Activation by Bacterial Tyrosyl-tRNA Synthetase. *J. Mol. Biol.*, **346**, 105-117.
- Kobayashi, T., Sakamoto, K., Takimura, T., Sekine, R., Kelly, V.P., Kamata, K., Nishimura, S. et Yokoyama, S. (2005b) Structural basis of nonnatural amino acid recognition by an engineered aminoacyl-tRNA synthetase for genetic code expansion. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 1366-1371.
- Kogelnik, A.M., Lott, M.T., Brown, M.D., Navathe, S.B. et Wallace, D.C. (1998) MITOMAP: a human mitochondrial genome database -1998 update. *Nucleic Acids Res.*, **26**, 112-115.
- Korencic, D., Ahel, I., Schelert, J., Sacher, M., Ruan, B., Stathopoulos, C., Blum, P., Ibba, M. et Söll, D. (2004) A freestanding proofreading domain is required for protein synthesis quality control in Archaea. *Proc. Natl. Acad. Sci. U.S.A.*, **101**, 10260-10265.
- Korge, P. et Weiss, J.N. (2006) Redox regulation of endogenous substrate oxidation by cardiac mitochondria. *Am. J. Physiol. Heart Circ. Physiol.*, **291**, H1436-1445.
- Kumazawa, Y., Himeno, H., Miura, K.-I. et Watanabe, K. (1991) Unilateral aminoacylation specificity between bovine mitochondria and eubacteria. *J. Biochem.*, **109**, 421-427.
- Kumazawa, Y., Yokogawa, T., Miura, K. et Watanabe, K. (1988) Bovine mitochondrial tRNA<sup>Phe</sup>, tRNA<sup>SerAGY</sup> and tRNA<sup>SerUCN</sup>: preparation using a new detection method and their properties in aminoacylation. *Nucleic Acids Symp. Ser.*, **19**, 97-100.
- Kunst, C.B., Mezey, E., Brownstein, M.J. et Patterson, D. (1997) Mutations in SOD1 associated with amyotrophic lateral sclerosis cause novel protein interactions. *Nat. Genet.*, **15**, 91-94.
- Kuratani, M., Sakai, H., Takahashi, M., Yanagisawa, T., Kobayashi, T., Murayama, K., Chen, L., Liu, Z.-J., Wang, B.-C., Kuroishi, C., Kuramitsu, S., Terada, T., Bessho, Y., Shirouzu, M., Sekine, S.-I. et Yokoyama, S. (2006) Crystal Structures of Tyrosyl-tRNA Synthetases from Archaea. *J. Mol. Biol.*, **355**, 395-408.
- Kyrpides, N., Overbeek, R. et Ouzounis, C. (1999) Universal protein families and the functional content of the last universal common ancestor. *J. Mol. Evol.*, **49**, 413-423.

## L

- Labouesse, M. (1990) The yeast mitochondrial leucyl-tRNA synthetase is a splicing factor for the excision of several group I introns. *Mol. Gen. Genet.*, **224**, 209-221.
- Laskowsky, R.A., McArthur, M.W., Moss, D.S. et Thornton, J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.*, **26**, 283-291.
- Lee, J.W., Beebe, K., Nangle, L.A., Jang, J., Longo-Guess, C.M., Cook, S.A., Davisson, M.T., Sundberg, J.P., Schimmel, P. et Ackerman, S.L. (2006) Editing-defective tRNA synthetase causes protein misfolding and neurodegeneration. *Nature*, **443**, 50-55.
- Lee, N., Bessho, Y., Wei, K., Szostak, J.W. et Suga, H. (2000) Ribozyme-catalyzed tRNA aminoacylation. *Nature Struct. Biol.*, **7**, 28-33.
- Lengyel, P. (1966) Problems in protein biosynthesis. *J. Gen. Physiol.*, **49**, 305-330.
- Lenhard, B., Orellana, O., Ibba, M. et Weygand-Durasevic, I. (1999) tRNA recognition and evolution of determinants in seryl-tRNA synthesis. *Nucleic Acids Res.*, **27**, 721-729.
- Levengood, J., Sandro, F.A., Roy, H. et Ibba, M. (2004) Divergence in non-cognate amino acid recognition between class I and class II lysyl-tRNA synthetases. *J. Biol. Chem.*, **279**, 17707-17714.
- Levinger, L., Jacobs, O. et James, M. (2001) In vitro 3' end endonucleolytic processing defect in a human mitochondrial tRNA<sup>SerUCN</sup> precursor with the U7445C substitution, which causes non-syndromic deafness. *Nucleic Acids Res.*, **29**, 4334-4340.
- Levinger, L., Mörl, M. et Florentz, C. (2004) Mitochondrial tRNA 3' end metabolism and human disease. *Nucleic Acids Res.*, **32**, 5430-5441.
- Limbach, P.A., Crain, P.F. et McCloskey, J.A. (1994) Summary: The modified nucleosides of RNA. *Nucleic Acids Res.*, **22**, 2183-2196.
- Ling, J., Roy, H. et Ibba, M. (2007) Mechanism of tRNA-dependent editing in translational quality control. *Proc. Natl. Acad. Sci. U.S.A.*, **104**, 72-77.
- Lipman, R.S., Chen, J., Evilia, C., Vitseva, O. et Hou, Y.M. (2003) Association of an aminoacyl-tRNA synthetase with a putative metabolic protein in archaea. *Biochemistry*, **42**, 7487-7496.
- Liu, M. et Spremulli, L.L. (2000) Interaction of mammalian mitochondrial ribosomes with the inner membrane. *J. Biol. Chem.*, **275**, 29400-29406.
- Liu, W., Brock, A., Chen, S., Chen, S. et Schultz, P.G. (2007) Genetic incorporation of unnatural amino acids into proteins in mammalian cells. *Nat Methods*, **4**, 239-244.
- Loftfield, R.B. (1972) The mechanism of aminoacylation of transfer RNA. *Prog. Nucleic Acid Res. Mol. Biol.*, **12**, 87-128.
- Lovato, M.A., Chihade, J.W. et Schimmel, P. (2001) Translocation within the acceptor helix of a major tRNA identity determinant. *EMBO J.*, **20**, 4846-4853.

Lue, S.W. et Kelley, S.O. (2005) An aminoacyl-tRNA synthetase with a defunct editing site. *Biochemistry*, **44**, 3010-3016.

Lynch, D.C. et Attardi, G. (1976) Amino acid specificity of the transfer RNA species coded for by HeLa cell mitochondrial DNA. *J. Mol. Biol.*, **102**, 125-141.

## M

Maréchal, L., Guillemaut, P. et Weil, J.H. (1985) Sequences of two bean mitochondria tRNAs<sup>Tyr</sup> which differ in the level of post-transcriptional modification and have a prokaryotic-like large extra-loop. *Plant Mol. Biol.*, **5**, 347-351.

Margulis, L. (1981) *Symbiosis in cell evolution*. Freeman, W.H., San Francisco.

Marinus, M.G., Morris, N.R., Söll, D. et Kwong, T.C. (1975) Isolation and partial characterization of three *Escherichia coli* mutants with altered transfer ribonucleic acid methylases. *J. Bacteriol.*, **122**, 257-265.

Marquet, R., Isel, C., Ehresmann, C. et Ehresmann, B. (1995) tRNAs as primer of reverse transcriptases. *Biochimie*, **77**, 113-124.

Mathews, M.B. et Bernstein, R.M. (1983) Myositis autoantibody inhibits histidyl-tRNA synthetase: A model for autoimmunity. *Nature*, **304**, 177-179.

Mathews, M.B., Reichlin, M., Hughes, G.R. et Bernstein, R.M. (1984) Anti-threonyl-tRNA synthetase, a second myositis-related autoantibody. *J. Exp. Med.*, **160**, 420-434.

McClain, W.H. (1995) The tRNA identity problem: past, present and future. In Söll, D. et RajBhandary, U.L. (eds.), *tRNA: Structure, Biosynthesis, and Function*. Am. Soc. Microbiol. Press, Washington DC, pp. 335-347.

Milligan, J.F., Groebe, D.R., Witherell, G.W. et Uhlenbeck, O.C. (1987) Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA template. *Nucleic Acids Res.*, **15**, 8783-8798.

Möller, W. et Janssen, G.M.C. (1990) Transfer RNAs for primordial amino acids contain remnants of a primitive code at position 3 to 5. *Biochimie*, **72**, 361-368.

Mudge, S.J., Williams, J.H., Eyre, H.J., Sutherland, G.R., Cowan, P.J. et Power, D.A. (1998) Complex organisation of the 5'-end of the human glycine tRNA synthetase gene. *Gene*, **209**, 45-50.

Murakami, H., Ohta, A., Ashigai, H. et Suga, H. (2006) A highly flexible tRNA acylation method for non-natural polypeptide synthesis. *Nat. Methods*, **3**, 357-359.

Muramatsu, T., Nishikawa, K., Nemoto, F., Kuchino, Y., Nishimura, S., Miyazawa, T. et Yokoyama, S. (1988) Codon and amino-acid specificities of a transfer RNA are both converted by a single post-transcriptional modification. *Nature*, **336**, 179-181.

## N

- Nagel, G.M. et Doolittle, R.F. (1991) Evolution and relatedness in two aminoacyl-tRNA synthetase families. *Proc. Natl. Acad. Sci. U.S.A.*, **88**, 8121-8125.
- Nakanishi, K. et Nureki, O. (2005) Recent Progress of Structural Biology of tRNA Processing and Modification. *Mol. Cells*, **19**, 157-166.
- Nakatsu, T., Kato, H. et Oda, J. (1998) Crystal structure of asparagine synthetase reveals a close evolutionary relationship to class II aminoacyl-tRNA synthetase. *Nature Struct. Biol.*, **5**, 15-19.
- Normanly, J., Kleina, L.G., Masson, J.M., Abelson, J. et Miller, J.H. (1990) Construction of *Escherichia coli* amber suppressor tRNA genes. III. Determination of tRNA specificity. *J. Mol. Biol.*, **213**, 719-726.
- Normanly, J., Ogden, R.C., Horvath, S.J. et Abelson, J. (1986) Changing the identity of a transfer RNA. *Nature*, **321**, 213-219.
- Notredame, C., Higgins, D.G. et Heringa, J. (2000) T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.*, **302**, 205-217.

## O

- O'Donoghue, P. et Luthey-Schulten, Z. (2003) On the evolution of structure in aminoacyl-tRNA synthetases. *Microbiol. Mol. Biol. Rev.*, **67**, 550-573.
- O'Hanlon, T.P., Raben, N. et Miller, F.W. (1995) A novel gene oriented in a head-to-head configuration with the human histidyl-tRNA synthetase (HRS) gene encodes an mRNA that predicts a polypeptide homologous to HRS. *Biochem. Biophys. Res. Commun.*, **210**, 556-566.
- O'Sullivan, O., Suhre, K., Abergel, C., Higgins, D.G. et Notredame, C. (2004) 3DCoffee: combining protein sequences and structures within multiple sequence alignments. *J. Mol. Biol.*, **340**, 385-395.
- Ochsner, U.A., Sun, X., Jarvis, T., Critchley, I. et Janjic, N. (2007) Aminoacyl-tRNA synthetases: essential and still promising targets for new anti-infective agents. *Expert. Opin. Investig. Drugs*, **16**, 573-593.
- Otani, A., Slike, B.M., Dorrell, M.I., Hood, J., Kinder, K., Ewalt, K.L., Cheresch, D., Schimmel, P. et Friedlander, M. (2002) A fragment of human TrpRS as a potent antagonist of ocular angiogenesis. *Proc. Natl. Acad. Sci. U.S.A.*, **99**, 178-183.
- Otwinowski, Z. et Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.*, **276**, 307-326.

## P

- Park, S.G., Ewalt, K.L. et Kim, S. (2005a) Functional expansion of aminoacyl-tRNA synthetases and their interacting factors: new perspectives on housekeepers. *Trends Biochem. Sci.*, **30**, 569-574.
- Park, S.G., Kim, H.J., Min, Y.H., Choi, E.-C., Shin, Y.K., Park, B.-J., Lee, S.W. et Kim, S. (2005b) Human lysyl-tRNA synthetase is secreted to trigger proinflammatory response. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 6356-6361.
- Paukstelis, P.J., Coon, R., Madabusi, L., Nowakowski, J., Monzingo, A., Robertus, J. et Lambowitz, A.M. (2005) A tyrosyl-tRNA synthetase adapted to function in group I intron splicing by acquiring a new RNA binding surface. *Molecular Cell*, **17**, 417-428.
- Pauling, L. (1957) The probability of errors in the process of synthesis of protein molecules. *Festschrift für Prof. Dr. Arthur Stoll*, **Basel: Birkhauser Verlag**, pp 597-602.
- Perret, V., Florentz, C., Puglisi, J.D. et Giegé, R. (1992) Effect of conformational features on the aminoacylation of tRNAs and consequences on the permutation of tRNA specificities. *J. Mol. Biol.*, **226**, 323-333.
- Perret, V., Garcia, A., Grosjean, H., Ebel, J.-P., Florentz, C. et Giegé, R. (1990) Relaxation of transfer RNA specificity by removal of modified nucleotides. *Nature*, **344**, 787-789.
- Pezo, V., Metzgar, D., Hendrickson, T.L., Waas, W.F., Hazebrouck, S., Döring, V., Marlière, P., Schimmel, P. et Crécy-Lagard, V.D. (2004) Artificially ambiguous genetic code confers growth yield advantage. *Proc. Natl. Acad. Sci.*, **101**, 8593-8597.
- Pham, Y., Li, L., Kim, A., Erdogan, O., Weinreb, V., Butterfoss, G.L., Kuhlman, B. et Carter, C.W., Jr. (2007) A minimal TrpRS catalytic domain supports sense/antisense ancestry of class I and II aminoacyl-tRNA synthetases. *Mol. Cell*, **25**, 851-862.
- Praetorius-Ibba, M., Hausmann, C.D., Paras, M., Rogers, T.E. et Ibba, M. (2007) Functional association between three archaeal aminoacyl-tRNA synthetases. *J. Biol. Chem.*, **282**, 3680-3687.
- Praetorius-Ibba, M., Rogers, T.E., Samson, R., Kelman, Z. et Ibba, M. (2005) Association between Archaeal prolyl- and leucyl-tRNA synthetases enhances tRNA<sup>Pro</sup> aminoacylation. *J. Biol. Chem.*, **280**, 26099-26104.
- Price, S.R., Ito, N., Oubridge, C., Avis, M. et Nagai, K. (1995) Crystallization of RNA-protein complexes. I. Methods for large-scale preparations of RNA suitable for crystallographic studies. *J. Mol. Biol.*, **249**, 398-408.
- Pulkes, T., Siddiqui, A., Morgan-Hughes, J.A. et Hanna, M.G. (2000) A novel mutation in the mitochondrial tRNA<sup>Tyr</sup> gene associated with exercise intolerance. *Neurology*, **55**, 1210-1212.
- Putney, S.D. et Schimmel, P. (1981) An aminoacyl-tRNA synthetase bind to a specific DNA sequence and regulates its gene transcription. *Nature*, **291**, 632-635.



Putzer, H., Grunberg-Manago, M. et Springer, M. (1995) Bacterial Aminoacyl-tRNA synthetases: Genes and regulation of expression. In Söll, D. et RajBhandary, U.L. (eds.), *tRNA: Structure, Biosynthesis, and Function*. Am. Soc. Microbiol. Press, Washington, DC, pp. 293-333.

## Q

Qiu, X., Janson, C.A., Smith, W.W., Green, S.M., McDevitt, P., Johanson, K., Carter, P., Hibbs, M., Lewis, C., Chalker, A., Fosberry, A., Lalonde, J., Berge, J., Brown, P., Houge-Frydrych, C.S.V. et Jarvest, R.L. (2001) Crystal structure of *Staphylococcus aureus* tyrosyl-tRNA synthetase in complex with a class of potent and specific inhibitors. *Protein Sci.*, **10**, 2008-2016.

## R

Raffelsberger, T., Rossmannith, W., Thaller-Antlanger, H. et Bittner, R.E. (2001) CPEO associated with a single nucleotide deletion in the mitochondrial tRNA<sup>Tyr</sup> gene. *Neurology*, **57**, 2298-2301.

Ramaswamy, K., Saito, H., Murakami, H., Shiba, K. et Suga, H. (2004) Designer ribozymes: programming the tRNA specificity into flexizyme. *J. Am. Chem. Soc.*, **126**, 11454-11455.

Ramirez, B.L., Howard, O.M.Z., Dong, H.F., Edamatsu, T., Gao, P., Hartlein, M. et Kron, M. (2006) *Brugia malayi* Asparaginyl-Transfer RNA Synthetase Induces Chemotaxis of Human Leukocytes and Activates G-Protein-Coupled Receptors CXCR1 and CXCR2. *J. Infect. Dis.*, **193**, 1164-1171.

Randau, L., Munch, R., Hohn, M.J., Jahn, D. et Soll, D. (2005) Nanoarchaeum equitans creates functional tRNAs from separate genes for their 5'- and 3'-halves. *Nature*, **433**, 537-541.

Ribas de Pouplana, L.R., Frugier, M., Quinn, C.L. et Schimmel, P. (1996) Evidence that two present-day components needed for the genetic code appeared after nucleated cells separated from eubacteria. *Proc. Natl. Acad. Sci. U.S.A.*, **93**, 166-170.

Ribas de Pouplana, L. (2005) Question: Why does the genetic code includes only 20 amino acids. *IUBMB Life*, **57**, 523-524.

Ribas de Pouplana, L. et Schimmel, P. (2001) Two classes of tRNA synthetases suggested by sterically compatible dockings on tRNA acceptor stem. *Cell*, **104**, 191-193.

Rich, A. et RajBhandary, U.L. (1976) Transfer RNA: molecular structure, sequence, and properties. *Annu. Rev. Biochem.*, **45**, 805-860.

Riddles, P.W., Blakeley, R.L. et Zerner, B. (1983) Reassessment of Ellman's reagent. *Methods Enzymol.*, **91**, 49-60.

Riès-Kautt, M. et Ducruix, A. (1992) Phase diagrams. In Ducruix, A. et Giegé, R. (eds.), *Crystallization of Nucleic Acids and Proteins: A Practical Approach*. IRL Press, Oxford, pp. 195-218.

- Robertus, J.D., Ladner, J.E., Finch, J.T., Rhodes, D., Brown, R.S., Clark, B.F.C. et Klug, A. (1974) Structure of yeast phenylalanine tRNA at 3 Å resolution. *Nature*, **250**, 546-551.
- Rocak, S., Landeka, I. et Weygand-Durasevic, I. (2002) Identifying Pex21p as a protein that specifically interacts with yeast seryl-tRNA synthetase. *FEMS Microbiol. Lett.*, **214**, 101-106.
- Rodin, S.N. et Ohno, S. (1995) Two types of aminoacyl-tRNA synthetases could be originally encoded by complementary strands of the same nucleic acid. *Orig. Life Evol. Biosph.*, **25**, 565-589.
- Rodin, S.N. et Rodin, A.S. (2006) Partitioning of aminoacyl-tRNA synthetases in two classes could have been encoded in a strand-symmetric RNA world. *DNA Cell Biol.*, **25**, 617-626.
- Rohrer, S. et Berger-Bachi, B. (2003) FemABX peptidyl transferases: a link between branched-chain cell wall peptide formation and beta-lactam resistance in gram-positive cocci. *Antimicrob. Agents Chemother.*, **47**, 837-846.
- Romby, P. et Springer, M. (2003) Bacterial translational control at atomic resolution. *Trends Genet.*, **19**, 155-161.
- Rossmann, M.G. et Arnold, E. (2001) *Crystallography of biological macromolecules* Springer.
- Rossmann, M.G., Moras, D. et Olsen, K.W. (1974) Chemical and biological evolution of a nucleotide-binding protein. *Nature*, **250**, 194-199.
- Roy, H., Ling, J., Alfonzo, J. et Ibba, M. (2005) Loss of editing activity during the evolution of mitochondrial phenylalanyl-tRNA synthetase. *J. Biol. Chem.*, **280**, 38186-38192.
- Ruff, M., Krishnaswamy, S., Boeglin, M., Poterszman, A., Mitschler, A., Podjarny, A., Rees, B., Thierry, J.C. et Moras, D. (1991) Class II aminoacyl transfer RNA synthetases: Crystal structure of yeast aspartyl-tRNA synthetase complexed with tRNA<sup>Asp</sup>. *Science*, **252**, 1682-1689.
- Ryckelynck, M., Masquida, B., Giegé, R. et Frugier, M. (2005) An Intricate RNA Structure with two tRNA-derived Motifs Directs Complex Formation between Yeast Aspartyl-tRNA Synthetase and its mRNA. *J. Mol. Biol.*, **354**, 614-629

## S

- Sahashi, K., Yoneda, M., Ohno, K., Tanaka, M., Ibi, T. et Sahashi, K. (2001) Functional characterisation of mitochondrial tRNA<sup>Tyr</sup> mutation (5877-->GA) associated with familial chronic progressive external ophthalmoplegia. *J. Med. Genet.*, **38**, 703-705.
- Sampath, P., Mazumder, B., Seshadri, V., Gerber, C.A., Chavatte, L., Kinter, M., Ting, S.M., Digman, J.D., Kim, S., Driscoll, D.M. et Fox, P.L. (2004) Noncanonical function of glutamyl-prolyl-tRNA synthetase: gene-specific silencing of translation. *Cell*, **119**, 195-208.

- Sampson, J.R. et Uhlenbeck, O.C. (1988) Biochemical and physical characterization of an unmodified yeast phenylalanine transfer RNA transcribed *in vitro*. *Proc. Natl. Acad. Sci. U.S.A.*, **85**, 1033-1037.
- Scaglia, F., Vogel, H., Hawkins, E.P., Vladutiu, G.D., Liu, L.L. et Wong, L. (2003) Novel homoplasmic mutation in the mitochondrial tRNA<sup>Tyr</sup> gene associated with atypical mitochondrial cytopathy presenting with focal segmental glomerulosclerosis. *Am. J. Med. Genet.*, **123A**, 172-178.
- Scheper, G.C., van der Klok, T., van Anandel, R.J., van Berkel, C.G., Sissler, M., Smet, J., Muravina, T.I., Serkov, S.V., Uziel, G., Bugiani, M., Schiffmann, R., Krageloh-Mann, I., Smeitink, J.A., Florentz, C., Van Coster, R., Pronk, J.C. et van der Knaap, M.S. (2007) Mitochondrial aspartyl-tRNA synthetase deficiency causes leukoencephalopathy with brain stem and spinal cord involvement and lactate elevation. *Nat. Genet.*, **39**, 534-539.
- Schimmel, P., Giegé, R., Moras, D. et Yokoyama, S. (1993) An operational RNA code for amino acids and possible relationship to genetic code. *Proc. Natl. Acad. Sci. U.S.A.*, **90**, 8763-8768.
- Schimmel, P. et Pouplana, L.R.D. (2000) Footprints of aminoacyl-tRNA synthetases are everywhere. *Trends Biochem. Sci.*, **25**, 207-209.
- Schnare, M.N., Heinonen, T.Y., Young, P.G. et Gray, M.W. (1985) Phenylalanine and tyrosine transfer RNAs encoded by *Tetrahymena pyriformis* mitochondrial DNA: primary sequence, post-transcriptional modifications, and gene localization. *Curr. Genet.*, **9**, 389-393.
- Schön, A., Krupp, A., Gough, S., Berry-Lowe, S., Kannangara, G. et Söll, D. (1986) The tRNA required in the first step of chlorophyll biosynthesis is a chloroplast glutamate tRNA. *Nature*, **322**, 281-284.
- Seburn, K.L., Nangle, L.A., Cox, G.A., Schimmel, P. et Burgess, R.W. (2006) An active dominant mutation of glycyl-tRNA synthetase causes neuropathy in a Charcot-Marie-Tooth 2D mouse model. *Neuron*, **51**, 715--726.
- Senger, B., Auxilien, S., Englisch, U., Cramer, F. et Fasiolo, F. (1997) The modified wobble base Inosine in yeast tRNA<sup>Ile</sup> is a positive determinant for aminoacylation by isoleucyl-tRNA synthetase. *Biochemistry*, **36**, 8269-8275.
- Shiba, K., Schimmel, P., Motegi, H. et Noda, T. (1994a) Human glycyl-tRNA synthetase. Wide divergence of primary structure from bacterial counterpart and species-specific aminoacylation. *J. Biol. Chem.*, **269**, 30049-30055.
- Shiba, K., Suzuki, N., Shigesada, K., Namba, Y., Schimmel, P. et Noda, T. (1994b) Human cytoplasmic isoleucyl-tRNA synthetase: Selective divergence of the anticodon-binding domain and acquisition of a new structural unit. *Proc. Natl. Acad. Sci. U.S.A.*, **91**, 7435-7439.
- Shimada, N., Suzuki, T. et Watanabe, K. (2001) Dual mode of recognition of two isoacceptor tRNAs by mammalian mitochondrial seryl-tRNA synthetase. *J. Biol. Chem.*, **276**, 46770-46778.

- Sibler, A.P., Dirheimer, G. et Martin, R.P. (1983) The primary structure of yeast mitochondrial tyrosine tRNA. *FEBS Lett.*, **152**, 153-156.
- Silvian, L.F., Wang, J. et Steitz, T.A. (1999) Insights into editing from an ile-tRNA synthetase structure with tRNA<sup>Ile</sup> and mupirocin. *Science*, **285**, 1074-1077.
- Simos, G., Sauer, A., Fasiolo, F. et Hurt, E.C. (1998) A conserved domain within Arc1p delivers tRNA to aminoacyl-tRNA synthetases. *Mol. Cell*, **1**, 235-242.
- Sisido, M., Ninomiya, K., Ohtsuki, T. et Hohsaka, T. (2005) Four-base codon/anticodon strategy and non-enzymatic aminoacylation for protein engineering with non-natural amino acids. *Methods*, **36**, 270-278.
- Sissler, M., Delorme, C., Bond, J., Ehrlich, S.D., Renault, P. et Francklyn, C. (1999) An aminoacyl-tRNA synthetase paralog with a catalytic role in histidine biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.*, **96**, 8985-8990.
- Sohm, B., Frugier, M., Brulé, H., Olszak, K., Przykorska, A. et Florentz, C. (2003) Towards understanding human mitochondrial leucine aminoacylation identity. *J. Mol. Biol.*, **328**, 995-1010.
- Söll, D. (1993) Transfer RNA - an RNA for all seasons. In Gesteland, R. et Atkins, J. (eds.), *The RNA World*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, pp. 157-184.
- Soutourina, J., Blanquet, S. et Plateau, P. (2000) D-tyrosyl-tRNA<sup>Tyr</sup> metabolism in *Saccharomyces cerevisiae*. *J. Biol. Chem.*, **275**, 11626-11630.
- Spencer, A., Heck, A., Takeuchi, N., Watanabe, K. et Spremulli, L. (2004) Characterization of the human mitochondrial methionyl-tRNA synthetase. *Biochemistry*, **43**, 9743-9754.
- Springer, M., Graffe, M., Butler, J.S. et Grunberg-Manago, M. (1986) Genetic definition of the translational operator of the threonine tRNA ligase gene in *Escherichia coli*. *Proc. Natl. Acad. Sci.*, **83**, 4384-4388.
- Sprinzi, M. et Vassilenko, K.S. (2005) Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res.*, **33**, D139-D140.
- Starzyk, R.M., Webster, T.A. et Schimmel, P. (1987) Evidence for dispensable sequences inserted into a nucleotide fold. *Science*, **237**, 1614-1618.
- Steinberg, S. et Cedergren, R. (1994) Structural compensation in atypical mitochondrial tRNAs. *Nature Struct. Biol.*, **1**, 507-510.
- Studier, F.W. et Moffatt, B.A. (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.*, **189**, 113-130.
- Suzuki, H., Ueda, T., Taguchi, H. et Takeuchi, N. (2007) Chaperone properties of mammalian mitochondrial translation elongation factor Tu. *J. Biol. Chem.*, **282**, 4076-4084.
- Suzuki, T., Suzuki, T., Wada, T., Saigo, K. et Watanabe, K. (2002) Taurine as a constituent of mitochondrial tRNAs: new insights into the functions of taurine and human mitochondrial diseases. *EMBO J.*, **21**, 6581-6589.

# T

- Tamura, K., Dudley, J., Nei, M. et Kumar, S. (2007) MEGA4 : Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4.0. *Mol. Biol. Evol.*, in press.
- Tanaka, T. et Yo, K. (2001) Origin of the cloverleaf shape of transfer RNA - the double hairpin model: Implication for the role of tRNA intron and the long extra loop. *Viva Origino*, **29**, 134-142.
- Targoff, I.N. (1990) Autoantibodies to aminoacyl-tRNA synthetases for isoleucine and glycine: Two additional synthetases are antigenic in myositis. *J. Immunol.*, **144**, 1737-1743.
- Théobald-Dietrich, A., Frugier, M., Giegé, R. et Rudinger-Thirion, J. (2004) Atypical archaeal tRNA pyrrolysine transcript behaves towards EF-Tu as a typical elongator tRNA. *Nucleic Acids Res.*, **32**, 1091-1096.
- Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. et Higgins, D.G. (1997) The CLUSTAL\_X windows interface: flexible strategies for multiple alignment by quality analysis tools. *Nucleic Acids Res.*, **25**, 4876-4882.
- Tolkunova, E., Park, H., Xia, J., King, M.P. et Davidson, E. (2000) The human lysyl-tRNA synthetase gene encodes both the cytoplasmic and mitochondrial enzymes by means of an unusual splicing of the primary transcript. *J. Biol. Chem.*, **275**, 35063-35069.
- Torres-Larios, A., Dock-Bregeon, A.-C., Romby, P., Rees, B., Sankaranarayanan, R., Caillet, J., Springer, M., Ehresmann, C., Ehresmann, B. et Moras, D. (2002) Structural basis of translational control by *Escherichia coli* threonyl-tRNA synthetase. *Nature Struct. Biol.*, **9**, 343-347.
- Tsunoda, M., Kusakabe, Y., Tanaka, N., Ohno, S., Nakamura, M., Senda, T., Moriguchi, T., Asai, N., Sekine, M. Yokogawa, T., Nishikawa, K. et Nakamura, K.T. (2007) Structural basis for recognition of cognate tRNA by tyrosyl-tRNA synthetase from three kingdoms. *Nucleic Acids Res.*, in press.
- Trifonov, E. et Berezovsky, I. (2002) Molecular evolution from abiotic scratch. *FEBS Lett.*, **527**, 1-4.
- Tumbula, D.L., Becker, H.D., Chang, W.Z. et Söll, D. (2000) Domain-specific recruitment of amide amino acids for protein synthesis. *Nature*, **407**, 106-110.
- Turner, J., Graziano, J., Spraggon, G. et Schultz, P. (2005) Structural Characterization of a p-Acetylphenylalanyl Aminoacyl-tRNA Synthetase. *J. Am. Chem. Soc.*, **127**, 14976-14977.
- Turner, J.M., Graziano, J., Spraggon, G. et Schultz, P.G. (2006) Structural plasticity of an aminoacyl-tRNA synthetase active site. *Proc. Natl. Acad. Sci. U.S.A.*
- Tzima, E., Reader, J.S., Ewalt, K.L., Schwartz, M.A. et Schimmel, P. (2003) Biologically active fragment of a human tRNA synthetase inhibits fluid shear stress-activated responses of endothelial cells. *Proc. Natl. Acad. Sci. U.S.A.*, **100**, 14903-14907.

Tzima, E. et Schimmel, P. (2005) Inhibition of tumor angiogenesis by a natural fragment of a tRNA synthetase. *Trends Biochem. Sci.*

## W

Wakasugi, K. et Schimmel, P. (1999) Two distinct cytokines released from a human aminoacyl-tRNA synthetase. *Science*, **284**, 147-150.

Wakasugi, K., Slike, B.M., Hood, J., Ewalt, K.L., Cheresch, D.A. et Schimmel, P. (2002) Induction of angiogenesis by a fragment of human tyrosyl-tRNA synthetase. *J. Biol. Chem.*, **277**, 20124-20126.

Wang, L., Xie, J. et Schultz, P.G. (2006) Expanding the Genetic Code. *Annu. Rev. Biophys. Biomol. Struct.*, **35**, 225-249.

Whelan, S. et Goldman, N. (2001) A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol. Biol. Evol.*, **18**, 691-699.

Widmann, J., Di Giulio, M., Yarus, M. et Knight, R. (2005) tRNA creation by hairpin duplication. *J. Mol. Evol.*, **61**, 524-530.

Wittenhagen, L.M. et Kelley, S.O. (2003) Impact of disease-related mitochondrial mutations on tRNA structure and function. *Trends Biochem. Sci.*, **28**, 605-611.

Woese, C.R., Olsen, G.J., Ibba, M. et Söll, D. (2000) Aminoacyl-tRNA synthetases, the genetic code, and the evolutionary process. *Microbiol. and Mol. Biol. Reviews*, **64**, 202-236.

Wolf, Y.I., Aravind, L., Grishin, N.V. et Koonin, E.V. (1999) Evolution of aminoacyl-tRNA synthetases--analysis of unique domain architectures and phylogenetic trees reveals a complex history of horizontal gene transfer events. *Genome Res.*, **9**, 689-710.

Wolfe, C.L., Warrington, J.A., Davis, S., Green, S. et Norcum, M.T. (2003) Isolation and characterization of human nuclear and cytosolic multisynthetase complexes and the intracellular distribution of p43/EMAPII. *Protein Sci.*, **12**, 2282-2290.

Wolfe, C.L., Warrington, J.A., Treadwell, L. et Norcum, M.T. (2005) A three-dimensional working model of the multienzyme complex of aminoacyl-tRNA synthetases based on electron microscopic placements of tRNA and proteins. *J. Biol. Chem.*, **280**, 38870-38878.

Wong, J.T.-F. (2005) Coevolution theory of the genetic code at age thirty. *Bioessays*, **27**, 416-425.

## X

Xayaphoummine, A., Bucher, T. et Isambert, H. (2005) Kinefold web server for RNA/DNA folding path and structure prediction including pseudoknots and knots. *Nucleic Acids Res.*, **33**, W605-610.

Xie, J. et Schultz, P.G. (2006) A chemical toolkit for proteins--an expanded genetic code. *Nat. Rev. Mol. Cell. Biol.*, **7**, 775-782.

Xue, S., Calvin, K. et Li, H. (2006) RNA recognition and cleavage by a splicing endonuclease. *Science*, **312**, 906-910.

## Y

Yanagisawa, T., Ishii, R., Fukunaga, R., Nureki, O. et Yokoyama, S. (2006) Crystallization and preliminary X-ray crystallographic analysis of the catalytic domain of pyrrolysyl-tRNA synthetase from the methanogenic archaeon *Methanosarcina mazei*. *Acta Cryst.*, **F62**, 1031-1033.

Yang, X.-L., Otero, F.J., Skene, R.J., McRee, D.E., Schimmel, P. et de Pouplana, L.R. (2003) Crystal structures that suggest late development of genetic code components for differentiating aromatic side chains. *Proc. Natl. Acad. Sci. U.S.A.*, **100**, 15376-15380.

Yang, X.L., Skene, R.J., McRee, D.E. et Schimmel, P. (2002) Crystal structure of a human aminoacyl-tRNA synthetase cytosine. *Proc. Natl. Acad. Sci. U.S.A.*, **99**, 15369-15374.

Yanisch-Perron, C., Vieira, J. et Messing, J. (1985) Improved M13 phage cloning vectors and host strain: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene*, **33**, 103-119.

Yaremchuk, A., Krikiviyi, I., Tukalo, M. et Cusack, S. (2002) Class I tyrosyl-tRNA synthetase has a class II mode of cognate tRNA recognition. *EMBO J.*, **21**, 3829-3840.

Yasukawa, T., Suzuki, T., Ishii, N., Ohta, S. et Watanabe, K. (2001) Wobble modification defect in tRNA disturbs codon-anticodon interaction in a mitochondrial disease. *EMBO J.*, **20**, 4794-4802.

Yokogawa, T., Shimada, N., Takeuchi, N., Benkowski, L., Suzuki, T., Omori, A., Ueda, T., Nishikawa, K., Spremulli, L.L. et Watanabe, K. (2000) Characterization and tRNA recognition of mammalian mitochondrial seryl-tRNA synthetase. *J. Biol. Chem.*, **275**, 19913-19920.

## Z

Zhang, Y., Wang, L., Schultz, P.G. et Wilson, I.A. (2005) Crystal structures of apo wild-type *M. jannaschii* tyrosyl-tRNA synthetase (TyrRS) and an engineered TyrRS specific for O-methyl-L-tyrosine. *Protein Sci.*, **14**, 1340-1349.